

Article

A Demand Forecasting Strategy Based on a Retrofit Architecture for Remote Monitoring of Legacy Building Circuits

Rubens A. Fernandes ^{1,2,*} , Raimundo C. S. Gomes ^{1,2} , Carlos T. Costa, Jr. ² , Celso Carvalho ³ ,
Neilson L. Vilaça ^{1,2} , Lennon B. F. Nascimento ¹ , Fabricio R. Seppe ¹ , Israel G. Torné ¹ 
and Heitor L. N. da Silva ¹ 

- ¹ Embedded Systems Laboratory, State University of Amazonas, Manaus 69050-020, Brazil; rsgomes@uea.edu.br (R.C.S.G.); neilsonluniere@gmail.com (N.L.V.); lennonbfn@gmail.com (L.B.F.N.); fab.seppe@gmail.com (F.R.S.); itorne@uea.edu.br (I.G.T.); heitor.lns@outlook.com (H.L.N.d.S.)
- ² Programa de Pós-Graduação em Engenharia Elétrica—PPGEE, Federal University of Pará, Belém 66075-110, Brazil; cartav@ufpa.br
- ³ Departamento de Eletrônica e Computação—DTEC, Federal University of Amazonas, Manaus 69067-005, Brazil; ccarvalho_@ufam.edu.br
- * Correspondence: rubens.eng.elet@gmail.com; Tel.: +55-92-98212-9068

Abstract: Energy demand forecasting is crucial for planning and optimizing the use of energy resources in building facilities. However, integrating digital solutions and learning techniques into legacy buildings presents significant challenges due to limited or outdated resources, hampering predictive analytics in these buildings and their circuits. To fill this gap, this article proposes an innovative demand forecasting strategy using an AIoT retrofit architecture based on the SmartLVGrid metamodel. This architecture allows remote monitoring of legacy building circuits, facilitating the collection, processing and storage of data in the cloud. We use several learning algorithms, including linear regression, support vector regressor, random forest regressor, XGBoost regressor, and long short-term memory (LSTM) neural network, to predict energy demand 15 min ahead, identifying potential overruns of contracted demand in accordance with Brazilian regulations. After Bayesian optimization, the LSTM neural network outperformed other models for most of the selected datasets and detected 32 out of 38 demand overruns on the test set. XGBoost and random forest followed closely, detecting 30 demand overruns. Overall, our cost-effective solution optimizes energy usage and efficiently mitigates potential demand exceedances in building installations. This is achieved through a step-by-step approach to upgrading existing aging facilities, which promotes energy efficiency and sustainability.

Keywords: demand forecast; retrofit; SmartLVGrid; AIoT; machine learning; real-time energy monitoring; energy efficiency; sustainability; smart buildings



Citation: Fernandes, R.A.; Gomes, R.C.S.; Costa, C.T., Jr.; Carvalho, C.; Vilaça, N.L.; Nascimento, L.B.F.; Seppe, F.R.; Torné, I.G.; da Silva, H.L.N. A Demand Forecasting Strategy Based on a Retrofit Architecture for Remote Monitoring of Legacy Building Circuits. *Sustainability* **2023**, *15*, 11161. <https://doi.org/10.3390/su151411161>

Academic Editor: Tullio De Rubéis

Received: 13 June 2023
Revised: 7 July 2023
Accepted: 12 July 2023
Published: 18 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Digital paradigms, including internet of things (IoT), and smart buildings and cities, are enabling the efficient use of resources essential for daily activities, such as electricity and water. In addition, they help in better decision making regarding the management of these resources, promoting scalability, flexibility, and dynamism characterized by the so-called data-driven approach [1,2]. However, the digital transformation of legacy systems still presents challenges such as a lack of support and updates, incompatibilities, and insufficient resources to interact with current systems. Alternatively, updating these systems can occur through a process of gradual and less costly technological transformation compared to the complete replacement of legacy systems [3–5]. Thus, using strategies that promote the digital transformation of legacy infrastructures can be a viable alternative for acquiring data and information for data-driven management of legacy systems.

Despite maintaining a significant portion of its legacy resources, the electricity sector is essential for the development of numerous socioeconomic activities. This can be observed by

the correlation between the increase in energy demand and the modernization of society [6,7]. Energy demand is a fundamental parameter for issues such as sustainability and energy efficiency, as it subsidizes the dimensioning of energy resources to meet society's needs. However, most legacy systems do not have resources for monitoring or forecasting demand in real time, making it impossible to take actions to reduce or optimize energy demand. Additionally, the lack of these resources makes it impossible to forecast exceedances of the contracted demand of companies and industries with energy concessionaires, which may result in fines or increases in the energy tariff of building installations. Thus, the use of digital solutions to monitor and forecast energy demand represents an opportunity to upgrade and optimize legacy resources.

Artificial intelligence of things (AIoT) can enable the management of electricity in terms of decentralized remote monitoring and computational resources for demand forecasting or energy consumption prediction [8,9]. Nevertheless, the literature lacks demand forecasting strategies based on energy parameters of legacy systems, which in many cases require interoperability resources and real-time monitoring. Without these, accessing the accurate demand profile of existing facilities and their circuits becomes a challenge for forecasting tasks using statistical methods or learning models.

In this context, retrofitting can be a strategy to update existing systems with digital solutions, preserving their resources and infrastructure [10,11]. However, to perform retrofitting systematically, allowing flexibility, scalability, and standardized integration with legacy systems, a reference model with well-defined protocols and interfaces is required. The SmartLVGrid metamodel enables the digital convergence of electrical systems to the smart grids paradigm [3,12]. In the literature, this metamodel has been used to achieve smart building convergence in legacy buildings to promote energy efficiency through resources for managing energy demand and electrical parameters in building installations [4,5].

However, there is a gap in the state of the art regarding the use of statistical techniques and artificial intelligence to predict energy demand in legacy building circuits. In this sense, we propose a legacy circuit retrofitting architecture based on a reference model to monitor electrical circuits and generate a monitoring database that can be used to implement energy demand forecast models for the installation and its circuits. This allows for a systematic and non-abrupt strategy for modernizing existing resources, allowing demand management and forecasting in the operations of building facilities. Furthermore, this proposal may enable the implementation of the strategy in other cases and systems.

In this article, we proposed a demand forecasting strategy in legacy building systems based on the retrofitting of these facilities. In our proposal, we presented a retrofit architecture to integrate hardware devices into a building power distribution panel capable of collecting and transmitting real-time data to the cloud. These data were further processed using supervised learning techniques to predict the energy demand of both the facility and its circuits. We used the SmartLVGrid metamodel at the physical and architectural levels as a basis to retrofit the legacy installation, ensuring the necessary interfaces and interoperability between monitoring devices and the cloud application created for data storage and processing.

With the data acquired by the proposed monitoring system, we conducted an exploratory analysis of the consumption and demand data from the installation and its circuits to mitigate the potential exceedance of the contracted demand in the legacy building installation of this study, following the regulatory standards for energy supply and distribution in Brazil, where the proposal was validated. Consequently, we performed short-term demand forecasting for the next 15 min. As learning models, we employed the random forest regressor (RFR), support vector regression (SVR), XGBoost regressor (XGBR), and a long short-term memory (LSTM)-based neural network architecture. Additionally, we used the performance results of the linear regression (LR) model as a baseline for evaluating and comparing the performance metrics (root mean squared error—RMSE, mean absolute error—MAE, and R-squared score— R^2) obtained for the mentioned models.

Therefore, we highlight the following contributions of this work:

- (1) Developing an AIoT solution for energy demand forecasting in legacy buildings and their circuits based on a retrofit strategy;
- (2) Implementing and comparing the performance of demand forecasting models in legacy electrical circuits using different learning models;
- (3) Implementing a new real-time monitoring system for energy demand in legacy electrical circuits based on the SmartLVGrid metamodel;
- (4) Proposing a systematic method for creating databases through the monitoring of pre-existing circuits;
- (5) Developing an alternative for detecting exceedances of the contracted demand with energy utility companies in legacy building installations using learning models.

To present our proposal, we divide the paper as follows: Section 2 provides a survey of the state of the art related to the topic. In Section 3, we highlight the research gaps in the literature concerning the theme of this work. Section 4 provides the theoretical framework of the SmartLVGrid metamodel. Section 5 presents our proposal for energy monitoring based on retrofitting low-voltage legacy circuits of a power distribution panel. In Section 6, we define our strategy and methodology to enable demand forecasting in the building installation and its legacy circuits. Section 7 presents the obtained results. In Section 8, we discuss the results, followed by the conclusions and proposals for future work in Section 9.

2. Related Work

The forecasting of energy demand is constantly researched in the literature, as well as the prediction of energy consumption. Among the approaches used in this context, statistical methods, machine learning, or deep learning models can be mentioned, employed based on pre-established databases. The most commonly used statistical methods are based on autoregressive techniques, with the most common ones being autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) methods. In [13], the SARIMA method was used by the authors to predict energy consumption in Poland on a quarterly, monthly, and weekly scale, using data from 2015 to 2021. In [14], the authors used the ARIMA method to estimate energy demand in Brazil from 2021 to 2025 and evaluated the predictability of the model using real data from the period 2014 to 2015. The authors of [15] also employed the SARIMA method to forecast short-term energy consumption for the Brazilian industrial sector. These statistical methods have also been used in the literature to make predictions using time series by rearranging the data present in the datasets to enable the forecasting of future energy demand based on past demand values. In the works [16,17], the authors used the sliding window method and autoregressive models to enable predictions of short-term future demands.

Although statistical methods have shown significant results in time series forecasting, they are well-suited when the dataset exhibits well-defined seasonality and trend patterns. When the time series exhibits more complex and even nonlinear patterns, machine learning methods can provide better results compared to statistical methods [18]. In [19], the authors proposed models for predicting electricity consumption in Slovakia using artificial neural networks. The authors of [20] used the support vector regression (SVR) and generalized regression neural network (GRNN) models to predict energy consumption in Indonesia. In the work [21], the authors applied random forest regression (RFR) and SVR to predict medium-term electricity demand using a Canadian database. In [22], the authors applied two ensemble learning methods, the XGBoost regressor and RFR, to forecast demand for the next day during the pandemic period. In the work [23], the authors employed machine learning methods, including linear regression (LR), multivariate polynomial regression, SVR, gradient boosting regressor (GBR), RFR, and K-neighbors regressor, to predict energy demand in New South Wales, Australia. In [24], the authors developed a clustering-based method for electricity prediction that was evaluated using a dataset with data from 105 substations. In the work [25], the authors presented a summary of the works developed in the IEEE demand forecasting competition, which included anomalous consumption data from a metropolitan region during the COVID-19 pandemic period. Various data preprocessing and demand prediction methods

using machine learning were presented. In an analysis of the cited works, it is mentioned that in cases of large data volume, nonlinear relationships among the characteristics present in the database, the presence of noise, and non-stationary behaviors, deep neural networks can be an alternative to machine learning. However, it is emphasized that deep networks require more computational resources and are more complex compared to supervised machine learning models. It is also mentioned that authors commonly use recurrent neural networks in this scenario, especially LSTM networks, combined with sliding window techniques [26–31]. Tables 1 and 2 summarize the previously presented works.

Table 1. Studies employing statistical methods for demand and energy consumption prediction.

Work	Year	Application	Methods or Models	Dataset Origin
[13]	2021	Prediction of electricity consumption in Poland on a quarterly, monthly, and semi-annual scale.	XGBoost, GRNN, SARIMA, ETS, NNETAR	Cire.pl
[14]	2022	Forecast of Brazilian monthly energy demand.	RS, ES, ARIMA	ONS Brazil
[15]	2022	Prediction of monthly consumption of industrial electricity in the Brazilian energy system.	HW, SARIMA, TBATS, DLM, NNAR, MLP	Central Bank of Brazil
[16]	2022	Out-of-sample, monthly, weekly, and hourly forecast for Nord Pool electricity demand.	AR, FAR, FARX	Nord Pool
[17]	2022	Short-term forecast of hourly energy demand of different energy districts.	SLFN, ARIMA, SVR, LSTM	Arpae, ARPA Lombardia

The abbreviations are presented in the list of abbreviations.

Table 2. Machine and deep learning studies for demand and energy consumption prediction.

Work	Year	Application	Methods or Models	Dataset Origin
[19]	2022	Development of electricity forecasting models in Slovakia.	Gray Models, ANN	Damas (SEPS)
[20]	2022	Electricity prediction in Bali Island, located in Indonesia, using electricity and weather data.	SVR, GRNN	East Java Province, domestic generators, ERA5-ECMWF
[21]	2021	Use of machine and deep learning models for medium-term prediction in Canada.	LSTM, SVR, NARX, RFR	IESO (Canada), Gov. of Canada
[22]	2022	Forecast for the next day of energy demand in Germany in COVID-19 pandemic period.	Ensemble-based models	OPSD
[23]	2022	Prediction of energy demand in New South Wales, Australia.	LR, MPR, SVR, ENR, GBR, DTR, RFR, KNNR	AEMO, Gov. of Australia
[24]	2022	Energy prediction based on cluster optimization method.	Greedy clustering	Ausgrid
[25]	2022	Demand prediction works in a metropolitan region using machine learning, statistical methods, and hybrid models.	Ensemble methods, AR, LR	BluWave-ai
[26]	2022	Short-term energy forecast using learning models.	ARIMA, LSTM, Prophet, Hybrid models	Elia grid
[27]	2020	Long-term demand prediction in Florida with regression models.	MRM, CNN variants, RFR, LSTM	EIA (U.S.), FCC, Census Bureau (U.S.), Bureau of Labor Statistics (U.S.)
[28]	2022	Prediction of energy consumption in Spain using LSTM networks.	LSTM variants	Spain Electricity Consumption
[29]	2021	Use of LSTM and convolutional networks for short-term demand forecasting in France and Korea.	LSTM and CNN variants	UCI repository, local Korean dataset
[30]	2023	Forecasting energy consumption demand using TFT, which outperformed other deep learning models.	LSTM variants, TCN, TFT	London DataStore
[31]	2022	Energy consumption forecasting on smart grids with N-BEATS, outperforming other deep learning methods.	LSTM and GRU variants, TCN, N-BEATS	London DataStore

The abbreviations are presented in the list of abbreviations.

The previously cited works contribute to the state of the art in demand forecasting and energy consumption. However, these works focus on predictions and forecasts relevant to energy companies, regional, or national contexts, rather than being directly related to building and industrial facilities. Additionally, the datasets employed were not produced through wireless sensor networks (WSNs) developed and configured by the authors, which would allow for the investigation of specific details or aspects, such as the use of predictive models for energy demand control, for example.

Thus, we sought literature that investigates the building context and applications of demand forecasting specifically tailored to building installations. In [32], demand and generation prediction of renewable energy sources, specifically photovoltaic and wind energy, were conducted in five smart residences using LSTM networks as prediction models, with approximately 11 months of collected data. In [33], an energy management strategy based on demand classification and prediction was presented. In addition to predicting the demand for a commercial building in Singapore, the authors developed neural network algorithms for decision making regarding energy excess treatment, application of photovoltaic energy, and energy storage conditions in the battery bank. In [34], the authors used a FFANN model for demand forecasting in the next 24 h for residential, educational, and mixed-use buildings. The authors of [35] predicted energy consumption in a food company based on data obtained from the factory's energy management system using the SVR and multilayer perceptron (MLP) methods. The work in [36] presents a study to assist managers and technicians with long-term energy predictions for a building at Teesside University (UK) using different machine learning techniques such as SVR and neural networks. In [37], the authors performed demand prediction using LSTM networks applied to the context of smart buildings. In [38], energy consumption data from smart meters installed in building substations, which recorded the consumption of the entire building at 15-min intervals, were utilized. Based on this data, the authors analyzed the integration of methods for consumption forecasting to improve energy efficiency in building installations. Table 3 presents the works cited in this paragraph on demand forecasting and energy consumption in building and industrial infrastructures.

Table 3. Research on forecasting demand and consumption of electricity in building and industrial infrastructures.

Work	Year	Application	Methods or Models	Dataset Origin
[32]	2021	Energy prediction and for renewable sources in smart buildings.	LSTM variants	HUE dataset (Harvard dataverse)
[33]	2020	Prediction and classification of energy demand for decision making in smart buildings.	MLP, RNN, LSTM, GRU, EM-GMM, BGM, K-means	Own data
[34]	2019	Use of the FFANN model to forecast demand for the next 24 h of buildings.	FFANN	Buildings of Finland
[35]	2023	Energy prediction in a food company using machine learning models.	MLP and SVR variants	Own data, KEPCO, KMA
[36]	2021	Long-term energy prediction in a university building.	PR, SVR, ANN	Own data
[37]	2022	Prediction of energy demand in smart buildings.	ARIMA, LSTM	Mendeley data
[38]	2022	Forecasting of energy consumption in smart buildings with different drift detection methods.	RFR, XGBoost CNN, TCN	Own data

The abbreviations are presented in the list of abbreviations.

Additionally, we selected some works that incorporate the concept of AIoT for electrical energy analysis. In [39], the authors developed a hardware device to monitor human presence and energy consumption. By using a decision tree model on a cloud-stored database, they determined energy waste in residential consumer units. Using the same decision tree algorithm, the authors of [40] created an energy control system based on hardware with wifi communication, relays, current sensors, and cloud storage. In the

work [41], neural networks were employed to predict energy consumption based on data collected from sensors in a residential system. The authors utilized these predictions to turn off one or more devices to reduce monthly energy consumption. The authors of [42] addressed the challenges of thermal management in electric vehicle batteries and proposed an AIoT-based preventive diagnostic system to improve safe driving, efficient maintenance, and product lifecycle management, aiming to optimize efficiency and battery life. Table 4 summarizes the selected AIoT works.

Table 4. Literature works on AIoT implementation in energy applications.

Work	Year	Application	Methods or Models	Dataset Origin
[39]	2022	IoT solution to control consumption and energy waste in homes.	Decision tree	Own data
[40]	2021	AIoT solution for controlling energy consumption in smart homes.	Decision tree	Own data
[41]	2022	Use of neural networks to control energy consumption in homes from wireless sensor networks (WSNs).	ANN	Own data
[42]	2023	AIoT system for preventive diagnosis of thermal challenges in electric vehicle batteries.	ANN	Own data

The abbreviations are presented in the list of abbreviations.

3. Research Gap

Previous studies on demand and energy consumption forecasting have shown the potential to enhance energy efficiency in building and industrial infrastructures within their respective contexts. However, there are several gaps in the current state of the art regarding demand or energy consumption forecasting in building facilities:

- Most existing studies rely on databases generated by third parties, without real-time AIoT solutions specifically designed to construct databases that capture patterns or characteristics of not only the overall electrical installation but also individual circuits and sectors within it. This presents an opportunity to leverage demand or consumption forecasting algorithms to optimize operations for specific installations of interest;
- The studies have not explored the forecasting of energy consumption and demand at the circuit level within building installations, which would enable individual analysis of high-consumption loads within the facility. This limitation stems from the lack of digital monitoring solutions that can collect individual demand data from building circuits, in addition to capturing the overall energy demand of the facility;
- The existing works do not provide AIoT solutions that enable the forecasting or detection of demand exceedances in legacy building systems, hindering digital convergence in pre-existing environments. A sustainable technological alternative is needed to promote energy efficiency in these installations. Retrofit strategies could be employed to introduce computational resources and update legacy infrastructures, leveraging existing resources to extract consumption and energy demand data for specific studies focused on legacy installations;
- The studies do not utilize retrofit strategies or metamodels with generic architectures and protocol stacks to enable systematic data collection through digital solutions that incorporate control, monitoring, distributed processing, and communication capabilities within data networks. Such approaches would benefit various cases and applications in the domain of energy forecasting.

Therefore, this study proposes to address these gaps by developing and implementing digital solutions using retrofit techniques and the SmartLVGrid metamodel for accurate demand forecasting in legacy installations.

4. SmartLVGrid

A smart low-voltage grid, or SmartLVGrid, is a metamodel that enables the technological convergence of legacy power distribution systems into the smart grid paradigm through retrofit strategies and systems engineering concepts. Its proposal involves adding electronic and computational resources for the control and monitoring of legacy systems using supervisory systems hosted on a local network or even in the cloud. These functionalities are described in the platform as operational primitives (OPs), which were previously performed by field operators and later, with the implementation of the metamodel, taken over by the added technological resources. This metamodel consists of protocol stacks described in two layers: middleware and interoperability, as shown in Figure 1.

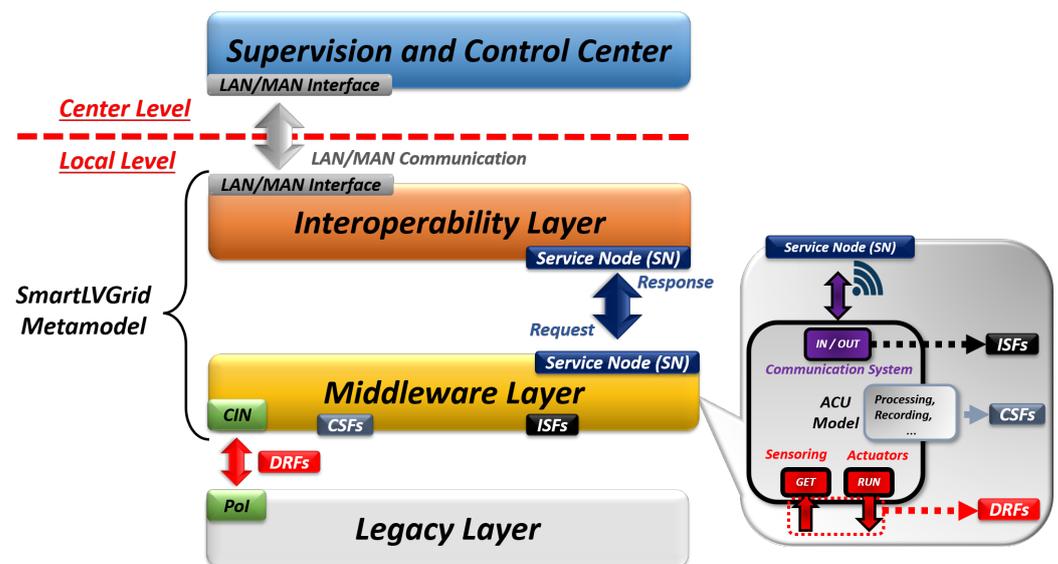


Figure 1. The SmartLVGrid stack [4].

As illustrated in Figure 1, the retrofitting of the existing infrastructure (legacy layer) is carried out through points of interface (PoIs) that interact with the middleware layer through the coupling and interaction node (CIN). Through this interface, the metamodel defines one of its operational primitives (OPs) called the domain retrofitting function (DRF), which is responsible for performing control and monitoring functions in the legacy layer. On the other hand, the service nodes (SNs) enable the middleware layer to interact with the interoperability layer through predefined communication standards and protocols. Thus, communication processes are performed by the interdomain support functions (ISFs). It should be noted that in the middleware layer, computational support functions (CSFs) are implemented to provide processing and storage services. In the following paragraphs and Sections 4.1 and 4.2, more details about the middleware and interoperability layers will be provided.

4.1. Middleware Layer

The middleware layer, which interacts directly with the legacy layer, is implemented through retrofitting solutions. Typically, these solutions encompass hardware devices with embedded processing, including sensor and actuator elements compatible with the DRFs to be executed. Alternatively, the middleware layer is described as the automation and communication unit (ACU), as shown in Figure 1. The ACU has “In/Out” ports that perform the communication processes, “Get” and “Run”, responsible for monitoring functionalities and controlling the legacy system, respectively. It should be noted that the CSFs are executed through the storage and processing resources of the ACU.

4.2. Interoperability Layer

The interoperability layer enables communication between ACUs through a data network. Additionally, the communication protocols and device hierarchies modeled through the SmartLVGrid metamodel are established within the interoperability layer. In this context, the ACUs that supervise and collect data from other ACUs, as well as execute DRFs when applicable, are hierarchically referred to as ACU coordinators. On the other hand, the supervised ACUs that execute DRFs in the legacy layer are called ACU operators. In cases of expanding the legacy system, it may be necessary to increase the computational capacity of the ACU coordinator. In the metamodel, it is possible to define sub-coordinators for each cluster of ACU operators, as described in [4]. Thus, sub-coordinators are associated with a single ACU coordinator, which transfers system information to and from the supervisory center. It is important to emphasize that, due to the local processing capability of each ACU, actions and directives can be performed by the ACU itself at the local level, enabling distributed and decentralized processing.

5. Methodology for Implementing the Energy Monitoring System

In previous works, we utilized wifi network infrastructures for communication with the supervisory centers [3–5]. However, in this study, we explore a different alternative for communication between our monitoring proposal and the supervisory center, as well as for the physical interface of the retrofit modules with the legacy building circuits, considering the specific characteristics of the monitored consumer unit. Specifically, we focus on a wifi router assembly factory where the main power distribution panel does not have sufficient space for installing retrofit modules, as shown in [5]. In this scenario, it is a factory regulation not to use wifi networks within its facilities to reduce interference issues and IP node conflicts during router testing and validation processes. Therefore, we employ a different retrofit approach compared to previous state-of-the-art works in terms of both physical and logical interfaces. Figure 2 illustrates the proposed retrofit strategy for the power distribution panel in the industry under study. Subsequently, Figure 3 presents an architecture diagram of the devices used in accordance with the SmartLVGrid metamodel, highlighting the adopted communication standards as well as the physical and logical interfaces of our monitoring proposal.

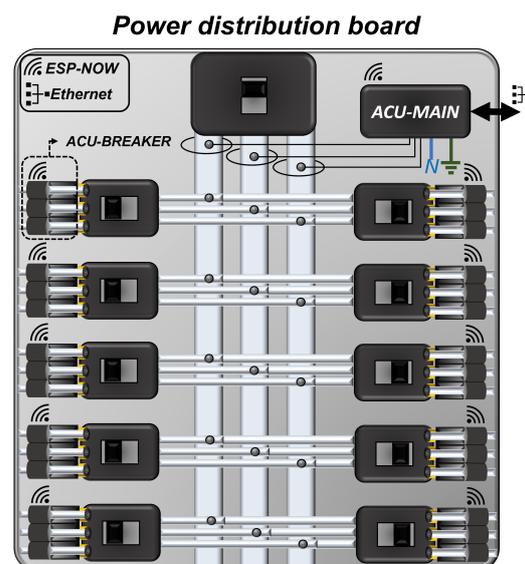


Figure 2. Retrofit strategy.

As depicted in Figure 2, the new strategy involves the integration of more compact retrofit modules compared to the modules developed in [5]. Still referred to as ACU-BREAKERS, in this study, the retrofit modules were powered by connecting them to the breakers of the main power distribution panel, enabling the monitoring of electrical

ACU-MAIN and the supervisory center was established through wired communication with a local server, adopting the MQTT protocol over ethernet. This allowed us to establish a connection with the cloud-hosted SCC. In summary, some benefits related to the hardware and communication architecture of our retrofit proposal include:

- Utilization of a peer-to-peer communication architecture among the wireless nodes, ACU-BREAKER (operator), and ACU-MAIN (coordinator), through the ESP-NOW ad hoc network, enabling communication flexibility and reducing the number of IP nodes;
- Adaptation of the monitoring modules, ACU-BREAKER, with a specific and compact design for installation in small-sized power distribution panels, reducing the space requirements and visual clutter of the industrial distribution panel;
- Development of retrofit modules that allow easy and intuitive installation in power distribution panels, thanks to the agile coupling features and reduced physical dimensions;
- Preservation of the existing resources in the installation, including the infrastructure, breakers, cables, connections, and the main distribution panel itself.

In this way, we enable the monitoring of the electrical panel and the forwarding of data to a local server for subsequent transmission to the cloud, where the supervisory and control center (SCC) is located. In the SCC, we built a dataset containing the obtained data from each circuit to be used in the demand prediction algorithms. Expanding its original proposal, the SCC now contributes not only with resources for storing and visualizing past information but also with predictive analysis resources for each circuit of the building installation through demand forecasting. The retrofit proposal tests were carried out by integrating and validating the physical integration and communication of the monitoring system with the cloud application, which receives the electrical parameters obtained from each circuit.

Subsequently, we present the modeling of the ACUs, compatible with the assumptions of the SmartLVGrid metamodel. The presented modeling will provide a detailed understanding of the conceived and developed physical and logical interfaces at the hardware and/or software level for the retrofit modules in the energy monitoring system.

5.1. ACU-BREAKER Conception and Modeling

Figure 4 presents the improved ACU-BREAKER (operator) developed during this work. The main differentiators of this ACU operator are its physical connection to the legacy circuits of the power distribution panel and the use of the ESP-NOW ad hoc protocol for communication between the ACU operators and the coordinator. As shown in the figure, it has metallic terminations that fit into the breakers and current transformers embedded in its structure. Therefore, the installation of the ACU-BREAKER is facilitated by inserting and screwing the connection cables of transformers/breakers onto the metallic terminations of the ACU-BREAKER. It is worth noting that the hardware and firmware resources and functionalities of the ACU-BREAKER are similar to those described in [5]. Thus, this ACU provides the DRF of electrical parameter monitoring through its Get port, performs ISFs of request and response through its In/Out port, and utilizes the ESP-NOW protocol for communication, along with CSFs related to network connection management, device configuration, and data storage. In terms of hardware, this device includes the same electronic surge protection devices, voltage and current channel conditioning, and ADE7758 for digitalization of acquired electrical parameters [47–49]. It is important to mention that the calibration procedures for the ACU-BREAKER, as described in [5], were maintained during the development of this work.

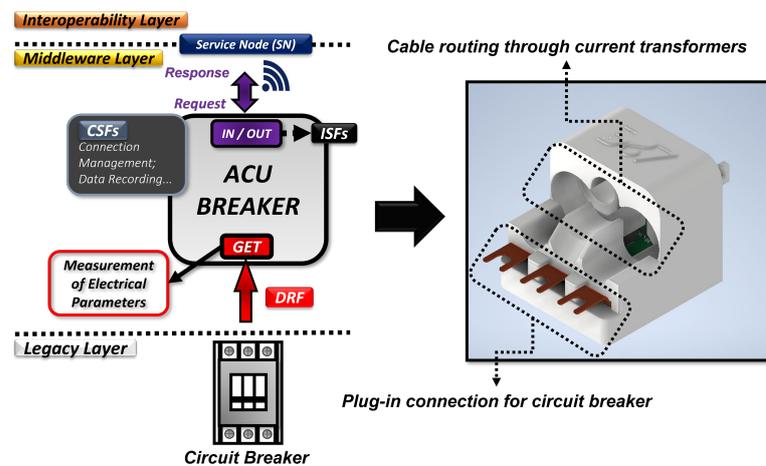


Figure 4. ACU-BREAKER architecture diagram and its physical perspective after development.

5.2. ACU-MAIN Conception and Modeling

The ACU-MAIN coordinator of the proposed system has similar DRFs, ISFs, and CSFs as the ACU-BREAKER. Additionally, it has the function of managing the network connection and communication with the other ACUs, including storing the identification data of the connected ACUs. Furthermore, it has an ethernet communication interface to communicate with the local server of the factory using the MQTT protocol [50–52]. The service nodes (SNs) of the SmartLVGrid metamodel for both the ACU-MAIN and ACU-BREAKER are established based on the credentials used in the ESP-NOW communication protocol, which includes the MAC address of the ESP32 used in the ACU hardware. It should be emphasized that the In/Out ports of this ACU are implemented through the ethernet interface for MQTT communication and the 2.4 GHz radio for ESP-NOW communication. The voltage and current parameters are monitored through the physical connection to the main bus and current transformers, respectively [53]. Figure 5 illustrates the ACU-MAIN developed in this work.

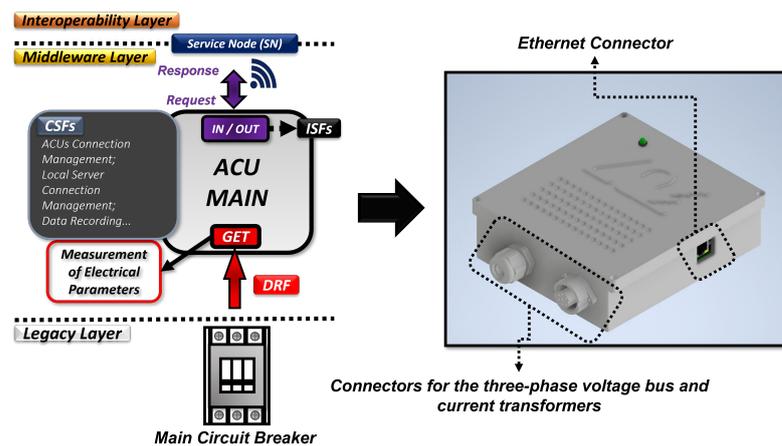


Figure 5. ACU-MAIN architecture diagram and its physical perspective after development.

5.3. Definition of the System Interoperability Layer

As mentioned earlier, the interoperability of the system occurs through two forms of communication. First, within the power distribution panel, the ACU operators communicate with the ACU-MAIN using the ESP-NOW wireless communication protocol. Second, the ACU-MAIN communicates with the local server of the factory through an ethernet interface, using the MQTT protocol with QoS 0. It should be noted that the ethernet interface was determined according to the company's requirements and aligns with the

retrofit concept of the SmartLVGrid metamodel, which aims to maximize the utilization of the existing legacy system. Consequently, the local server forwards the messages to an MQTT broker hosted on the DigitalOcean Droplet virtual server hosting service, also with QoS 0, where the processing of energy data takes place. It is important to mention that the request messages for electrical parameters are transmitted in JSON format and, upon receipt at the SCC, they are stored in a MongoDB database.

The service nodes (SNs), illustrated in Figures 4 and 5, represent the credentials that allow the ACUs to communicate in a wireless network. In this work, the SNs are implemented through the credentials that enable the communication of devices using the ESP-NOW protocol, including the MAC address of the ESP32 in each ACU in the proposed P2P interface.

Regarding the messages in our proposal, they are implemented using JSON format for both the interface between ACU operators and the ACU-MAIN and the interface between the ACU-MAIN and the local server. The same message protocol is also adopted for communication between the local server and the SCC. The messages include request and response messages for sending the monitored electrical parameters along with timestamps, network communication parameter changes, inclusion of new devices, and ACU-BREAKER calibration. Figure 6 illustrates the process adopted to enable the interoperability of our proposal in a request of electrical parameter scenario as follows:

- The local server requests the electrical parameters from the ACU operators and the ACU-MAIN every minute (1);
- The configuration of the service nodes (SNs) of the ACU-BREAKERS and the ACU-MAIN is performed (2);
- The request for electrical parameters is sent from the ACU-MAIN to each ACU-BREAKER using the ESP-NOW protocol (3);
- Upon receiving the request, the ACU-BREAKER performs ISFs to synchronize communication and transmits the requested data to the ACU-MAIN (4);
- After collecting the information from the ACUs and the message timestamps, the local server forwards the data to the cloud-hosted SCC (5).

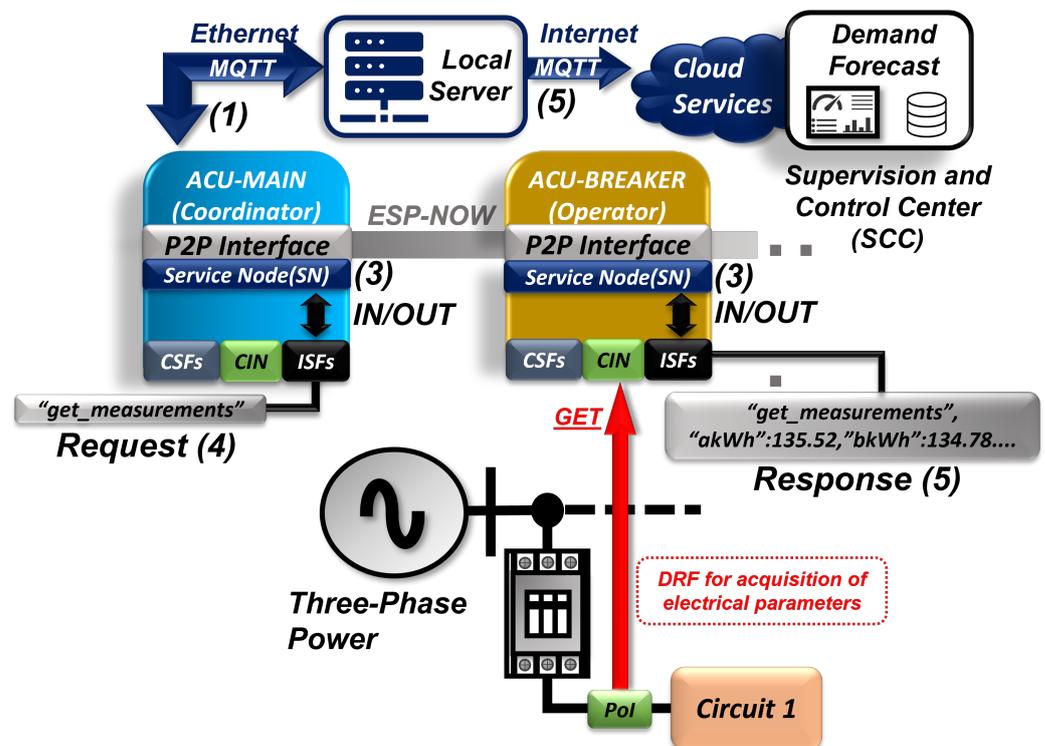


Figure 6. Communication process of the proposed system.

5.4. Installation of the ACUs

Once assembled, tested, and calibrated, the ACUs were installed and configured to operate in the existing power distribution panel of the router factory. Each ACU was calibrated beforehand to match the nominal currents and voltages of the breakers in the panel, with a maximum error of 1%, using a precision three-phase source and the internal registers of the ADE7758, the integrated circuit used in the ACUs for electrical parameter digitalization [54,55]. The panel operates with a phase-neutral voltage of $127 V_{rms}$ and has 22 circuits. Figure 7 illustrates the ACUs installed in the legacy power distribution panel. As depicted, the first distribution breaker does not have an ACU-BREAKER installed, as it was damaged during the evaluation period of the proposal.

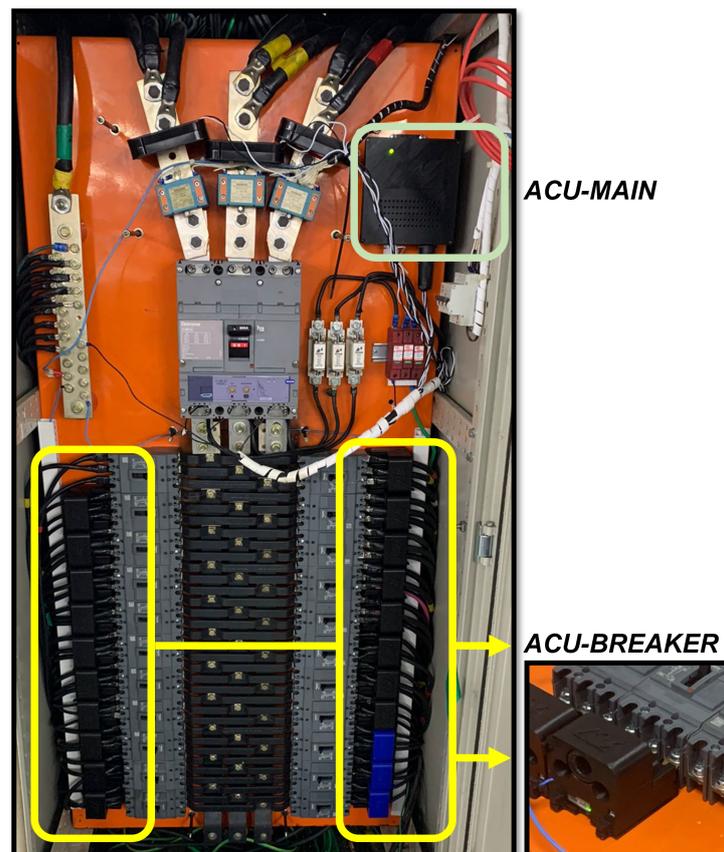


Figure 7. ACUs installed on legacy power distribution board.

6. Proposed Demand Forecast Strategy

The literature presents applications of the SmartLVGrid metamodel used for the management, control, and energy monitoring of power distribution systems and building systems [3,4,12]. In [5], we presented a data-driven energy management strategy by monitoring real-time energy demand in each circuit of a building installation based on the aforementioned metamodel. In Brazil, where the proposed work was implemented, medium- and high-voltage consumer units are categorized as “binomials”, being charged based on both consumption and previously contracted energy demand from a local energy distributor [56]. The demand is weighted every 15 min, and if it exceeds the stipulated value in the established contract, the consumer unit is subject to fines according to the Brazilian National Electric Energy Agency (ANEEL) in the normative resolution ANEEL No. 1000/2021 [57]. To assist the participating managers in the conducted case study, we also developed a visual interface with demand exceedance alarm indicators so that managers could choose to develop demand control strategies or renegotiate the demand contract with the energy distributor.

Thus, we noticed that a tool for predictive analysis of energy demand could contribute to anticipate potential exceedances and, if possible, act promptly to reduce costs associated with consumer demand exceeding limits, also assisting in demand management. Therefore, considering that each circuit in the legacy installation can be monitored through retrofit modules, the forecasting of demand for the next 15 min of the installation and its circuits could be performed at the supervision and control center (SCC), becoming an additional data analytics functionality incorporated into audit processes to enhance energy efficiency. Such a strategy would enable decision making for demand control or renegotiation of demand limits with the utility company, if necessary.

In this study, after installing the ACU-MAIN and ACU-BREAKERS in the main power distribution panel, we let the devices operate and collect individual data from each circuit, including the main breaker. The data were collected based on the interoperability definitions specified earlier in Section 5.3. The collected circuit parameters are detailed in Table 5. Subsequently, Table 6 presents the identification and load connected to each circuit, along with the monitoring system device that supervises the respective circuits.

Table 5. Data variable description.

Data Variable	Description
Circuit identification	Monitored circuit identification.
MAC address	MAC address of installed ACU.
Timestamp	Timestamp of samples (datetime format).
Power factor	Power Factor of each circuit (%).
Active energy	Active energy of each circuit (Wh).
RMS current	RMS current of each circuit (A).
RMS voltage	RMS voltage of each circuit (V).

Table 6. Circuit, load, and monitoring device description.

Circuit Identification	Load	Monitoring Device
Circuit 0	All Building Installation	ACU-MAIN
Circuit 2	Production Line—02	ACU-BREAKER-1
Circuit 3	Production Line—03	ACU-BREAKER-2
Circuit 4	Production Line—04	ACU-BREAKER-3
Circuit 5	Reserve Circuit	ACU-BREAKER-4
Circuit 6	Electrical Panel—Production	ACU-BREAKER-5
Circuit 7	Reserve Circuit	ACU-BREAKER-6
Circuit 8	Electrical Panel—Server 02	ACU-BREAKER-7
Circuit 9	Support Area—02	ACU-BREAKER-8
Circuit 10	Central Air Conditioning—01	ACU-BREAKER-9
Circuit 11	Support Area—03	ACU-BREAKER-10
Circuit 12	Administration	ACU-BREAKER-11
Circuit 13	Central Air Conditioning—02	ACU-BREAKER-12
Circuit 14	Electrical Panel—Stock 01	ACU-BREAKER-13
Circuit 15	Support Area—01	ACU-BREAKER-14
Circuit 16	Central Air Conditioning—03	ACU-BREAKER-15
Circuit 17	Electrical Panel—Stock 02	ACU-BREAKER-16
Circuit 18	Support Area—04	ACU-BREAKER-17
Circuit 19	Electrical Panel—Server 01	ACU-BREAKER-18
Circuit 20	Reserve Circuit	ACU-BREAKER-19
Circuit 21	Chamber	ACU-BREAKER-20
Circuit 22	Reserve Circuit	ACU-BREAKER-21

The proposed system transmits the collected data from minute to minute to the local server and then to the cloud. Based on this, it was possible to create a database at the SCC for conducting the study proposed in this work. The database used in this study was generated from 15 January to 12 April 2023, and contains data from the main breaker and 21 circuits of the distribution panel that supply loads and other distribution panels within

the building installation. Due to industrial confidentiality reasons, the obtained database and other company data could not be published or made available to the public at the moment, but we can make it available upon request and negotiations carried out directly with us. For the forecasting task proposed, only the minute-to-minute active energy data from each circuit will be used, which were subsequently processed to obtain the energy demand. The other data are used by the industry in energy audit procedures. It is important to mention that the building in question has a demand limit of 120 kW.

Throughout this section, we presented the exploratory analysis of the obtained data, the preprocessing techniques used for training the learning models, and the performance metrics for model evaluation. Hereafter, the concepts of the learning models used will be presented, followed by the division of the training and validation datasets.

In summary, to prepare the data for use in time series forecasting, we used the sliding window technique so that previous demand data could be used to predict future demand for the next 15 min for circuits within the installation, following the ANEEL guidelines in [57]. These data were normalized using the min–max method. Based on the performance of other works in the literature, we used machine learning regression techniques as learning models, such as random forest regressor (RFR), support vector regression (SVR), and XGBoost regressor (XGBR). Additionally, we used the linear regression (LR) method to obtain a prediction baseline from the preprocessed data, and a recurrent neural network model, specifically a long short-term memory (LSTM) network, as a deep learning alternative to compare with the other obtained results.

6.1. Exploratory Data Analysis and Definition of the Circuits to Be Analyzed

Before preprocessing the obtained data, we analyzed the contribution of each circuit to the energy consumption of the building installation. For this purpose, we performed a Pareto analysis of the total energy consumption of the circuits in the installation from 15 January to 12 April 2023. In this analysis, the cumulative percentage consumption was based on the ratio of the individual consumption of each circuit, monitored by the ACU operators, to the total consumption of the installation measured by the ACU-BREAKER. Circuit 0 represents the entire installation, which is monitored by the ACU-MAIN. The other circuits, from 2 to 22, are monitored through the ACU-BREAKERS. The Pareto diagram of the energy consumption of the circuits present in the installation is illustrated in Figure 8. It should be noted that, due to damage to the ACU-BREAKER of circuit 1 during the installation process and the fact that other circuits have much lower energy consumption compared to the rest, the total and percentage consumption of these circuits are identified as “other circuits” in the diagram.

We noticed that circuits 13, 16, 10, 8, 6, 12, and 14 accounted for approximately 80% of the total consumption of the installation. Since energy consumption is directly related to energy demand, we chose to perform demand forecasting studies for these circuits considering their contributions to the demand increase. In addition to these circuits, we also used the demand data obtained from the ACU-MAIN. From the energy data monitored every minute by the circuits, we extracted the 15-min energy demand for the mentioned circuits. Table 7 presents the statistical and descriptive data for 15-min demand intervals for the specified circuits. Here, “count” represents the number of demand values for each circuit’s dataset. Figure 9 illustrates box plots that detail the variation in the 15-min energy demand for these circuits.

We observed from Table 7 and Figure 9 that the average values of the 15-min demand are directly proportional to the cumulative percent of energy in Figure 8, justifying the selection of circuits based on Pareto analysis for the demand forecasting study. According to Table 7, the data count is the same for all samples collected from the selected circuits. From Table 7 and Figure 9, with the exception of circuits 6 and 8, we noticed that the largest deviations obtained are concentrated in the upper part of the graphs. We can observe from Table 7 that the standard deviation of the energy demand is more significant in the demand obtained from the monitored data of the main breaker of the distribution panel

(circuit 0). Additionally, it can be observed in Figure 9 that the graph indicates possible demand exceedances in the installation during the data collection period in this circuit, with values exceeding the contracted demand of 120 kW, as illustrated by the red marking in the figure. On the other hand, the outliers in the same figure are less frequent in the circuits of the main panel monitored by the ACU-BREAKERS. The circuits that present the most outliers are the demand data of circuits 8 and 12. We expect that the LSTM, SVR, RFR, and XGBR models perform better than the linear regression model in datasets with higher variability. The preprocessing techniques applied to the 15-min demand data, which are subsequently used in the training and testing of the learning models, will be presented next.

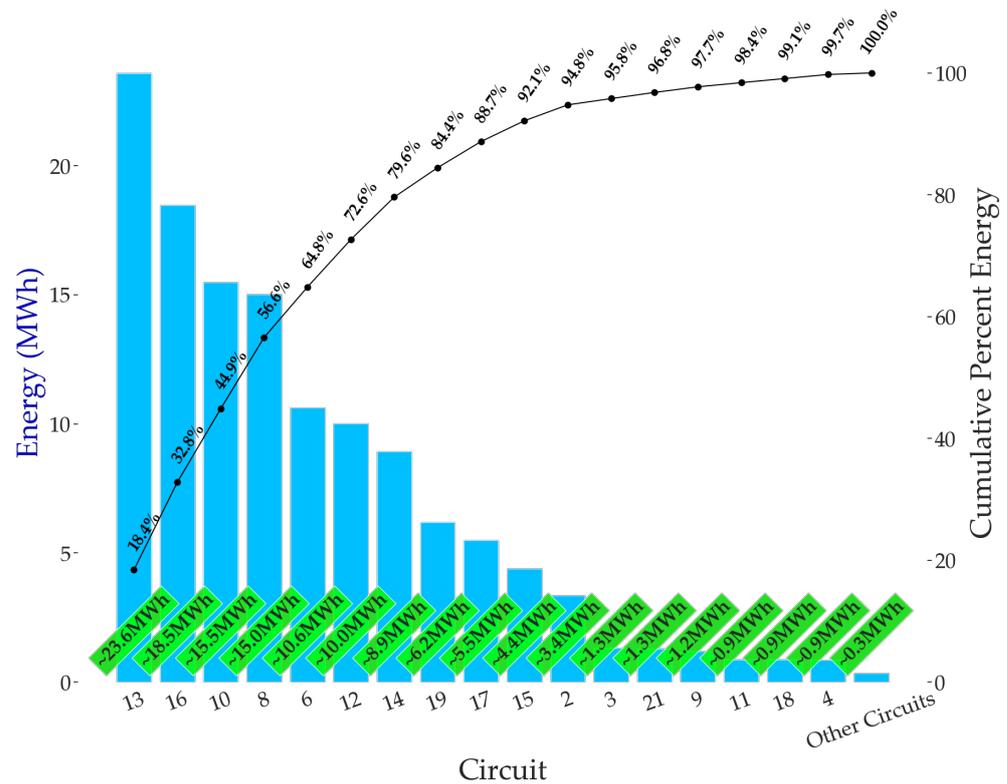


Figure 8. Pareto diagram of the energy consumption of building circuits.

Table 7. Descriptive statistics of the 15-min demand data.

Statistics	Circ. 0	Circ. 6	Circ. 8	Circ. 10	Circ. 12	Circ. 13	Circ. 14	Circ. 16
Count	6782	6782	6782	6782	6782	6782	6782	6782
Mean (kW)	62.18	6.27	8.86	9.13	5.90	13.92	5.27	10.90
Standard deviation (kW)	34.18	3.55	3.22	8.90	4.92	11.24	3.33	9.67
Lower value (kW)	9.25	0.54	0.09	0.11	0.66	0.46	0.45	0.11
First quartile (kW)	21.74	1.45	9.02	0.12	1.52	0.57	0.87	0.12
Median (kW)	72.35	8.57	9.79	5.49	3.76	20.14	6.51	11.11
Third quartile (kW)	89.60	8.79	10.44	19.52	9.80	24.33	8.13	22.30
Upper value (kW)	126.46	10.80	13.46	21.59	25.07	27.44	11.74	24.14

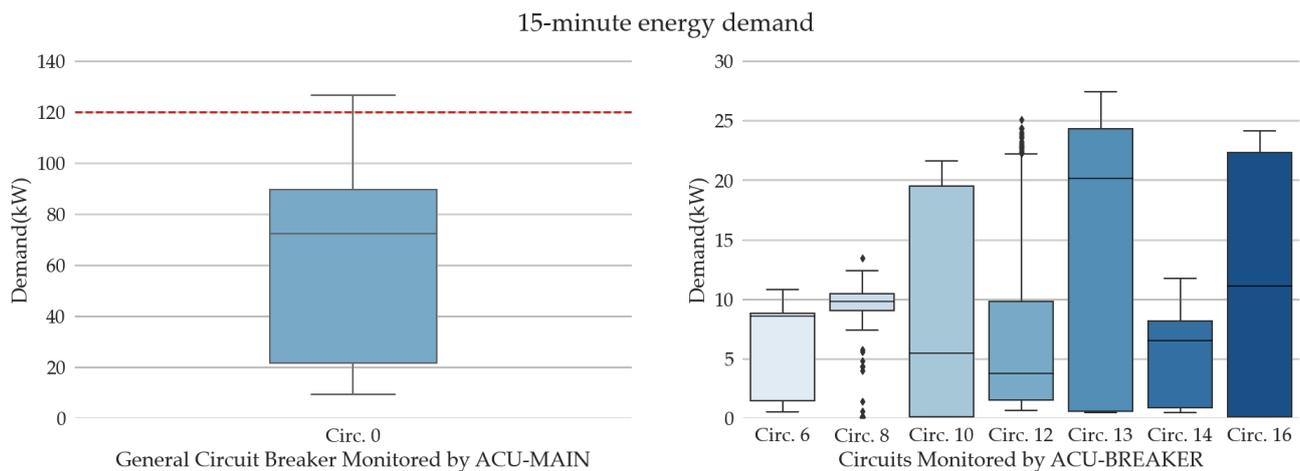


Figure 9. 15-min demand variation of the building installation and monitored circuits, with the contracted demand of the installation represented by a dashed red line.

6.2. Data Preprocessing

In this section, we present the methods used for data preprocessing in our study, which include the sliding window technique and min–max normalization. This crucial step ensures that the data entered into the models are in a suitable and ideal format for forecasting energy demand in the context of this work.

6.2.1. Sliding Window

The sliding window algorithm was used to generate the input data for the models by selecting subsets of sequential samples. These subsets are called sliding windows, which move with a predetermined temporal unit step according to each application [58]. This technique is widely used in areas such as time series forecasting, signal processing, and temporal data analysis. In this work, the temporal unit is defined as the energy demand values obtained from each circuit over a 15-min period. Each sliding window, as illustrated in Figure 10, is composed of past demand values (i.e., blue sets), which are used as input to predict the energy demand for the next temporal unit (i.e., cubes). We determined the optimal window size through empirical tests, where we established possible window values and performed iterative loops using the learning models. Based on the results obtained for each defined window, we have selected the best possible window size to predict the demand for the selected circuits. The window size determined from the conducted tests was 10 temporal units (samples) of 15 min of previous demands to predict the value of the energy demand for the subsequent sample.

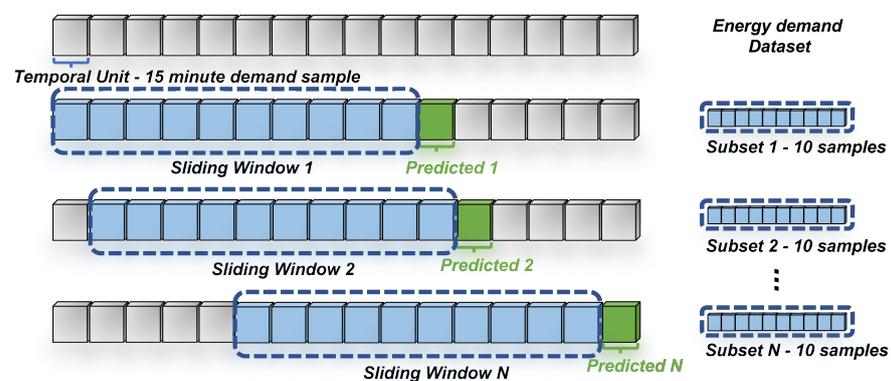


Figure 10. Sliding window technique.

6.2.2. Min–Max Normalization

The min–max data normalization method scales a dataset so that its values are within a specified range $[a, b]$. This technique is commonly used to preprocess data before applying machine learning algorithms. When applying min–max normalization to a dataset, the original values are transformed into new scaled values that fall within a specified range. This transformation is performed using an adaptation of the standard linear transformation, as shown in Equation (1). In this work, the range defined for data normalization was $[0, 1]$.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

6.3. Evaluation Metrics

In this section, we explain the critical metrics used to evaluate the performance of the implemented learning models. These evaluation metrics provide quantitative information about the performance of the models in forecasting energy demand.

6.3.1. Root Mean Squared Error—RMSE

Root mean squared error (RMSE) is a widely used metric for evaluating the performance of regression models. This measure assesses the difference between the actual values y_i and the predicted values \hat{y}_i of a dependent variable by calculating the square root of the mean of the squared errors, as shown in Equation (2).

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

By examining the equation of RMSE, it can be seen that the metric resembles the standard deviation. Thus, the RMSE value can be interpreted as a metric that indicates the variability in errors in relation to the actual values of the dependent variable. Therefore, it can be considered as an indicator of the model's accuracy, with a lower RMSE value indicating better performance. Additionally, the RMSE metric can be used as a quantitative measure of the prediction quality of the model for comparative analysis between regression techniques. It is worth noting the use of the square root, the RMSE can be interpreted in terms of the dependent variable, which helps in understanding the magnitude of errors generated by the evaluated model [59].

6.3.2. Mean Absolute Error—MAE

Mean absolute error (MAE) is an evaluation metric that provides the average magnitude of the n absolute differences between the predicted values y_i and the expected values \hat{y}_i . This metric is expressed in the same unit as the dependent variable and, therefore, provides a straightforward understanding and interpretation of the achieved performance, facilitating a direct comparison between different models [60]. The mathematical expression for MAE can be seen in Equation (3).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

6.3.3. R-Squared Score— R^2

The R-squared score (R^2) is an evaluation metric that indicates the proportion of the variance in the dependent/predicted variable y that is explained by the input/expected variables. This metric takes values between 0 and 1, where 0 indicates that the model does not explain any variability in the dependent variable, and 1 indicates that the model explains all the variability in the dependent variable. Therefore, as the R^2 value increases, the model fits the data better and explains a higher proportion of the variance in the dependent variable. On the other hand, an R^2 value close to 0 indicates that the model is

unable to explain the variation in the dependent variable [61]. This metric is expressed in Equation (4).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

6.4. Learning Models

In this section, we delve into the specificities of the learning models used in this work, which include linear regression, support vector regression, random forest regression (RFR), XGBoost regression, and LSTM-type recurrent networks.

6.4.1. Linear Regression (LR)

The linear regression (LR) method aims to establish a linear relationship between the response variable y and the predictor variables x_1, x_2, \dots, x_l , which are called the dependent and independent variables, respectively. In the context of demand prediction, the independent variable is the sampled data allocated in the window, while the dependent variable is the predicted demand. The linear relationship is obtained by estimating the parameter vector θ and adding an additive disturbance or noise term η . Thus, considering y_n as the demand at time n , and applying the sliding window, it follows that:

$$y_n = \theta_0 + \theta_1 y_{n-1} + \theta_2 y_{n-2} + \dots + \theta_l y_{n-l} + \eta_n \quad (5)$$

and

$$\eta_n = y_n - (\theta_0 + \theta_1 y_{n-1} + \theta_2 y_{n-2} + \dots + \theta_l y_{n-l}) \quad (6)$$

Considering N observations and $l = 10$, we have:

$$S(\theta) = \sum_{n=l+1}^N (y_n - \theta_1 y_{n-1} - \theta_2 y_{n-2} - \dots - \theta_l y_{n-l})^2 \quad (7)$$

$$S(\theta) = \sum_{n=l+1}^N \eta_n^2 \quad (8)$$

or in vector form:

$$S(\theta) = \sum_{n=l+1}^N (y_n - \theta^T \tilde{y}_n)^2 \quad (9)$$

where

$$\theta = (\theta_0, \theta_1, \dots, \theta_l)^T \quad (10)$$

and

$$\tilde{y}_n = (1, y_{n-1}, y_{n-2}, \dots, y_{n-l})^T \quad (11)$$

In this case, $w = (w_0, w_1, \dots, w_l)^T$ is the estimated vector of θ that minimizes $S(\theta)$. In general terms, the LR model performs a prediction by calculating the weighted sum of the input data and adding a constant term. This process determines the weights and biases of the model. In its multiple form, it involves the use of two or more predictors, i.e., more input variables for training. It is one of the most commonly used low-complexity models when the response variable and predictor have a strong linear correlation [62].

6.4.2. Support Vector Regression (SVR)

The SVR (support vector regression) prediction technique aims to predict output values by determining a hyperplane that closely resembles the input data. In this algorithm,

the maximum number of instances possible is considered within a margin of ϵ , with the aim of determining weights and biases, that provides the generalization for the model. To achieve this, the objective is to minimize the error $J(w, w_0, \zeta, \hat{\zeta})$ given by Equation (12), where ζ_n and $\hat{\zeta}_n$ are the slack variables corresponding to a deviation from the ϵ margin, with the penalty control given by C , constrained by Equations (13)–(15).

$$J(w, w_0, \zeta, \hat{\zeta}) = \frac{1}{2} \|w\|^2 + C \left(\sum_{n=1}^N \zeta_n + \sum_{n=1}^N \hat{\zeta}_n \right) \quad (12)$$

$$y_n - w^T x_n - w_0 \leq \epsilon + \hat{\zeta}_n, \quad n = 1, 2, \dots, N \quad (13)$$

$$w^T x_n + w_0 - y_n \leq \epsilon + \zeta_n, \quad n = 1, 2, \dots, N \quad (14)$$

$$\hat{\zeta}_n \geq 0, \zeta_n \geq 0, \quad n = 1, 2, \dots, N \quad (15)$$

In this way, contributions to the cost function from errors with an absolute value less than or equal to ζ are set to zero. The optimizer's objective is to estimate w and w_0 in a manner that the contribution of error values greater than ζ and smaller than $\hat{\zeta}$ is minimized. Thus, this algorithm is interesting for initial testing in machine learning and has the advantage of not being affected by local minima, unlike deep neural network algorithms. However, as the amount of data increases, this algorithm tends to lose performance when attempting to establish a linear response [63].

6.4.3. Random Forest Regression (RFR)

In a regression tree, the determination of the root node variable and subsequent nodes is defined by maximizing the weighted averages in the child nodes or, equivalently, by minimizing the weighted variance σ_w^2 of subsets Y_1, Y_2, \dots, Y_n , with $|Y_1|, |Y_2|, \dots, |Y_n|$ elements, as shown in Equation (16).

$$\sigma_w^2(Y_1, Y_2, \dots, Y_n) = \sum_{n=1}^N \frac{|Y_n|}{|Y|} \sigma^2(Y_n) \quad (16)$$

In the RF method, which is an algorithm based on an ensemble of decision trees, the bootstrap aggregating strategy is applied during the model learning phase. Bootstrap aggregating aims to construct a series of trees by randomly sampling the original data, using only a subset m of predictors from a complete set p of predictors. These samples are then trained independently and in parallel with each other. Finally, the values are aggregated by calculating the average of the results obtained from each individual regression tree [64].

Thus, by averaging multiple decision trees that are subjected to high variance, the model exhibits better generalization performance and is less prone to overfitting. The RF technique has been widely used to solve low-complexity regression problems due to its high performance and robustness against overfitting.

6.4.4. XGBoost Regressor (XGBR)

The XGBoost regressor algorithm is based on making predictions using regression decision trees. The method utilizes information aggregation, random forest for tree selection during batch training, error minimization using gradient descent, and regularization of weights and biases. Equations (17) and (18) present the weight function and the objective function, respectively. In these equations, g_i and h_i are the first- and second-order gradients of the loss function, λ and γ represent additional regularization terms, T represents the number of nodes, q represents the tree structure, and I_j is the instances of a node j . In addition to regularization, XGBoost uses an additional shrinkage technique to prevent overfitting by scaling the weights obtained by a factor η , similar to a learning rate. This

process reduces the influence of each individual tree and allows room for future trees to improve the model.

$$w_j = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (17)$$

$$J(q) = -\frac{1}{2} \sum_{j=1}^T \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (18)$$

This algorithm has shown promise in various prediction scenarios, including regression and classification problems. This is due to its high scalability, as the execution time of this algorithm can be 10 times faster than others, and it can be scaled for numerous examples in distributed configurations or with limited processing memory due to implemented optimizations and parallel processing capabilities [65].

6.4.5. Long Short-Term Memory (LSTM)

LSTM networks are a type of recurrent neural network that feature an internal memory cell structure as their main characteristic. Through the logistic function and multiplier weight matrices, these gates are implemented and referred to as the input gate (i_t), forget gate (f_t), and output gate (o_t). There is also the vector that represents the internal state (C_t) of the LSTM cell and the candidate value (\tilde{C}_t). The mathematical definitions of the gates, cell state, and candidate value of the LSTM network are presented in Equations (19)–(23), including the respective biases b_C , b_i , b_f , and b_o .

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (19)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (20)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (21)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \quad (22)$$

$$\tilde{C}_t \cong \tanh(W_c[h_{t-1}, x_t] + b_C) \quad (23)$$

The application of these networks is interesting for problems involving sequential data and time series, such as the electrical demand curve, for example, [66]. While a fully connected neural network has separate parameters for each input feature, recurrent neural networks share the same weights across different time steps, establishing a strong temporal relationship among the data.

6.5. Definition of Training and Test Sets

The demand data for the selected circuits consists of 6782 observations, as shown in Table 7. To proceed, we normalized the dataset using the min–max technique, we divided it into training and test subsets in order to implement and validate the learning models. Thus, 80% of the observations were used for training, and 20% were used for testing. Figure 11 illustrates the separated training and test sets for each circuit selected for the proposed demand prediction study in this work. After dividing the data, we applied the sliding window technique to prepare the input and output data subsets for training and testing the learning models. As mentioned earlier, the sliding window size adopted was 10 past values to predict a demand value for the next 15 min.

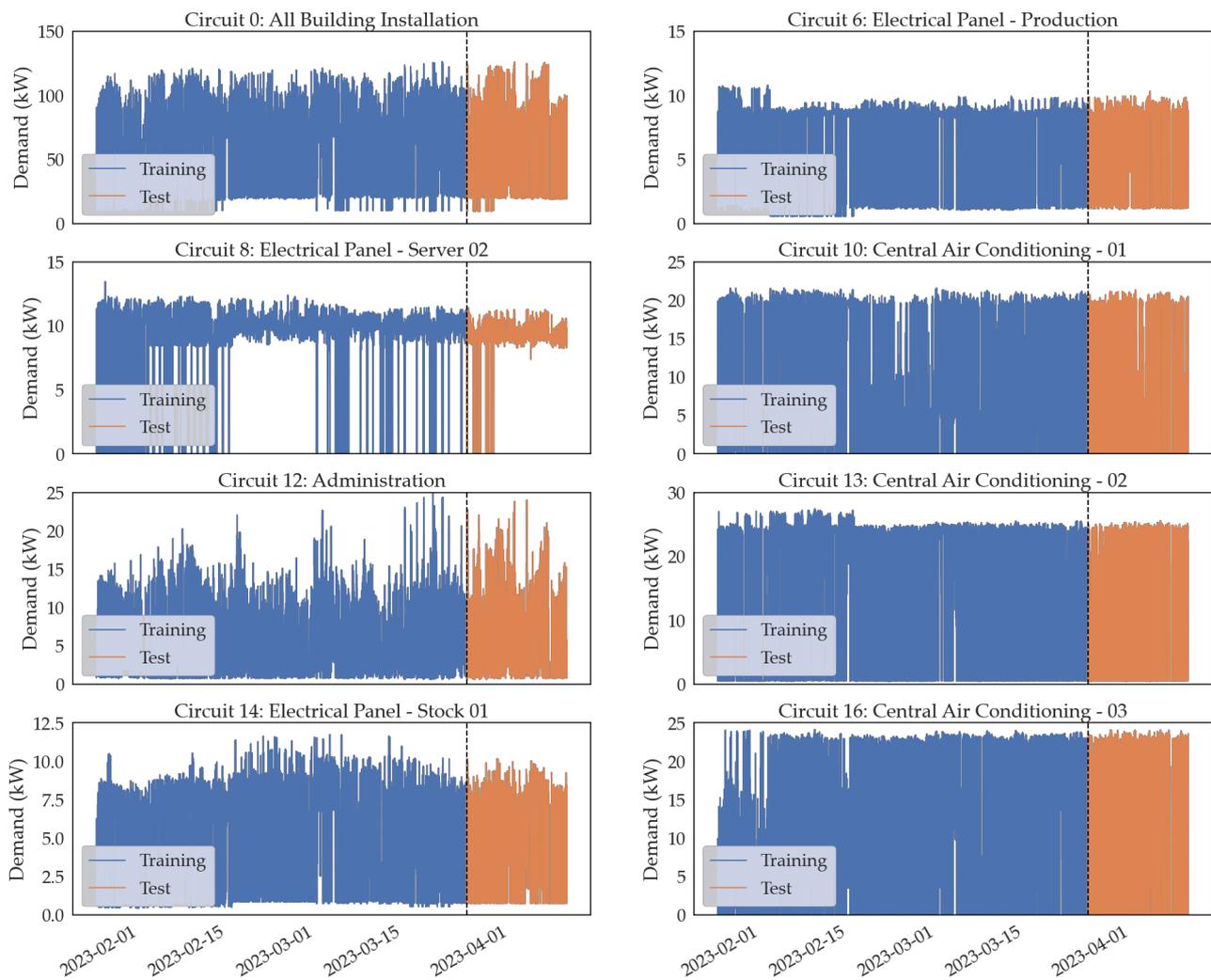


Figure 11. Training and test sets of the selected circuits.

The training of the models was carried out on a local server from the data collected in the SCC, where we evaluated the predictive models before transferring them back to the cloud server. The server has a 2.3 GHz Intel Core i7-11800H processor, 16 GB RAM, 4 GB GPU, and 500 GB SSD.

6.6. Software Libraries and Optimization of Learning Models

The experiments with the learning models were conducted on the Jupyter Lab platform of the Anaconda distribution using the Python language. We utilized several libraries, including TensorFlow, Pandas, NumPy, Matplotlib, Seaborn, XGBoost, and Scikit-learn. To enhance the performance of the learning models on the established dataset, we used the Optuna framework for Bayesian optimization of the hyperparameters of the machine learning models and fine-tuning of the LSTM model. Bayesian optimization techniques have proven to be more efficient in finding better hyperparameters and searching for the best parameters to be used in neural networks and their variants. This is because they make use of prior information about the behavior of the objective function to guide the search [67,68]. Optuna is an easy-to-configure Bayesian optimization framework that is suitable for hyperparameter tuning and determining the best parameters for supervised learning models for a given training and testing set. With a define-by-run API, the search space for the best parameters is dynamically defined by Optuna during the runtime of an objective function instantiated to test the desired model under pre-established conditions [69]. Thus, Optuna was used to train and evaluate

the models for each dataset of the selected circuits. The parameter K in the table represents the number of trees used in the RFR and XGBR models.

6.7. Definition of Parameters and Architectures of Learning Models

To accomplish the task of energy demand forecasting in our proposal, we conducted an investigation into various machine learning models to determine the most suitable one(s) for predicting the energy demand of the researched circuits, which exhibit distinct demand patterns. The architecture for evaluating the learning models is illustrated in Figure 12a, and the implemented LSTM model architecture is represented in Figure 12b. After conducting tests using the Optuna framework to evaluate the models, we were able to select the best parameters for each learning model. The tests were conducted individually for each model, considering the normalized datasets of circuits 0, 6, 8, 10, 12, 13, 14, and 16. We conducted 500 trials per study in an effort to find the optimal parameters that enabled the models to effectively capture the temporal demand characteristics. The mean squared error (MSE) metric was used as the evaluation criterion for training all the machine learning models. Table 8 showcases some of the hyperparameters discovered for the machine learning models after the Bayesian optimization process, considering the selected datasets.

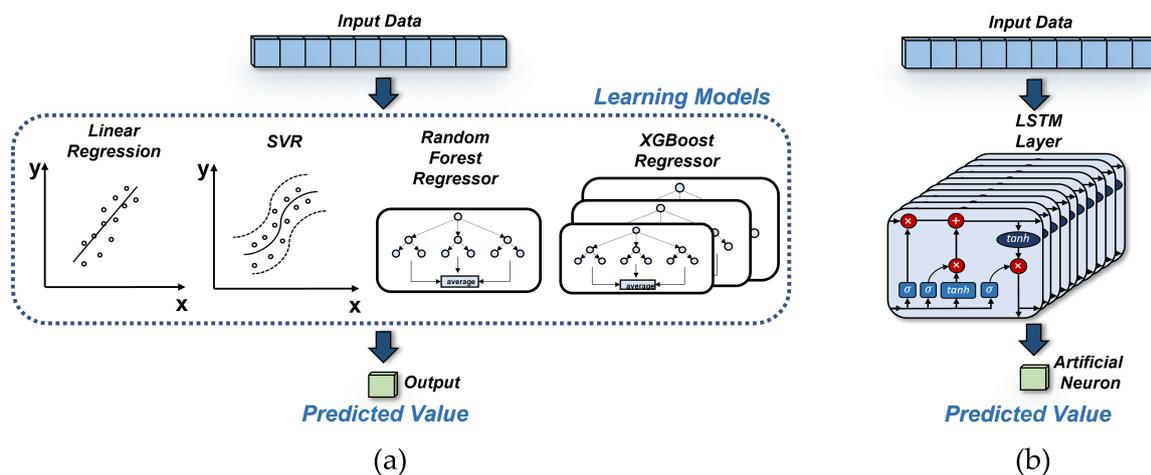


Figure 12. Learning models (a) and LSTM recurrent neural network model (b) used to evaluate demand forecasting.

Table 8. Hyperparameters used in machine learning models after optimization process.

Dataset	SVR	RFR	XGBR
Circ. 0	C: 115.495, ϵ : 0.011	K: 236	γ : 0.107, λ : 0.036, η : 0.207, K: 645
Circ. 6	C: 119.050, ϵ : 0.034	K: 558	γ : 0.273, λ : 0.898, η : 0.308, K: 530
Circ. 8	C: 53.516, ϵ : 0.011	K: 102	γ : 0.072, λ : 0.538, η : 0.242, K: 684
Circ. 10	C: 108.645, ϵ : 0.028	K: 498	γ : 0.407, λ : 0.238, η : 0.245, K: 505
Circ. 12	C: 51.044, ϵ : 0.014	K: 132	γ : 0.059, λ : 0.859, η : 0.260, K: 500
Circ. 13	C: 119.953, ϵ : 0.031	K: 217	γ : 0.254, λ : 0.284, η : 0.034, K: 555
Circ. 14	C: 108.214, ϵ : 0.018	K: 43	γ : 0.205, λ : 0.914, η : 0.246, K: 549
Circ. 16	C: 117.43, ϵ : 0.055	K: 408	γ : 0.255, λ : 0.458, η : 0.173, K: 569

K: number of trees.

When implementing the SVR, RFR, and XGBR models, it is crucial to understand the impact of the chosen parameters following the optimization process. In the case of SVR, the parameters C and ϵ control the regularization and error tolerance, respectively. Higher values of C can lead to overfitting, while very low values can result in underfitting. The parameter ϵ determines the width of the tolerance margin around the regression hyperplane. Therefore, the optimization process using the Optuna framework was crucial in selecting appropriate parameters and improving the SVR's performance. On the other hand, in the RFR model, the number of estimators (trees) K , determined through the optimization process, improves the model's generalization capability and reduces both the training and optimization times. The XGBR model also has several important parameters, such as the learning rate (η) and the number of estimators (K). The learning rate controls the contribution of each estimator in the update process. Lower values can lead to better generalization, while higher values can cause overfitting. The number of estimators affects the model's generalization capability and training time.

We also implemented an LSTM neural network model to compare with the LR, SVR, RFR, and XGBR models. In the implementation process of this model, we tested various architectures, including bidirectional LSTM networks and hybrid LSTM and convolutional networks. We also experimented with stacking LSTM layers to achieve better results. However, the best performance for the test set was obtained using a single LSTM layer with one artificial neuron in the output. We also utilized Optuna to optimize the parameters of the proposed LSTM network. Each Optuna trial for the LSTM network consisted of 100 training epochs using the Adam optimizer [66]. We conducted 500 trials for this model in the Optuna framework. The best parameters for this model are presented in Table 9. It is important to note that the activation function used in the LSTM layer of the models was the hyperbolic tangent (tanh).

Table 9. Best parameters for LSTM model on each dataset.

Dataset	Learning Rate	Units	Batch Size
Circ. 0	3.209×10^{-2}	38	70
Circ. 6	1.055×10^{-2}	23	24
Circ. 8	3.085×10^{-2}	20	23
Circ. 10	2.711×10^{-2}	80	24
Circ. 12	2.521×10^{-2}	80	70
Circ. 13	3.351×10^{-2}	48	64
Circ. 14	2.101×10^{-2}	28	36
Circ. 16	2.722×10^{-2}	80	64

The learning rate determines the step size used by the Adam optimization algorithm during the training of the LSTM. Low learning rates can result in slower convergence or become trapped in local minima, while high learning rates can make the training unstable and prevent the model from finding an optimal solution. The number of units determines the model's capacity to learn complex representations and capture patterns in the data. Higher values increase the learning capacity but also increase the training time and the need for more training data. The batch size determines the number of training samples used in each weight update pass of the LSTM. A larger batch size can speed up training by processing more samples in parallel. However, a larger batch size requires more memory, and training may become more challenging to parallelize. The choice of batch size depends on the available memory, the size of the training set, and the trade-off between training speed and accuracy. Thus, finding the appropriate parameters is crucial for striking a balance between training speed and the performance of the LSTM model.

7. Results

7.1. Performance Evaluation of Learning Models

Initially, we assessed the LR model's performance on the acquired datasets to establish a baseline for the performance metrics, to be achieved by the other learning models. After optimizing the learning models, we used the hyperparameters from Table 8 to evaluate the performance of the SVR, RFR, and XGBR models, and the parameters from Table 9 to evaluate the performance of the LSTM model. The performance metrics obtained for the learning models for the test subsets of each energy demand dataset are presented in Table 10. It is important to mention that the results presented for the performance metrics are not normalized, as the data were returned to their original scale after the models' predictions.

Table 10. Result of learning models' performance metrics for test sets of selected demands (non-normalized values).

Demand Dataset	RMSE (kW)					MAE (kW)					R ² (%)				
	LR	SVR	RFR	XGBR	LSTM	LR	SVR	RFR	XGBR	LSTM	LR	SVR	RFR	XGBR	LSTM
Circ. 0	9.116	8.789	8.269	8.252	8.216 *	4.874	4.278	4.152 *	4.273	4.285	92.705	93.22	93.998	94.02	94.07 *
Circ. 6	0.957	0.936	0.875	0.868	0.865 *	0.312	0.321	0.267	0.272	0.251 *	91.94	92.29	93.26	93.37	93.52 *
Circ. 8	0.426	0.417	0.424	0.420	0.415 *	0.215	0.199 *	0.217	0.214	0.205	86.81	87.35	86.90	87.16	87.39 *
Circ. 10	2.987	2.948	2.753	2.701 *	2.723	1.278	1.298	1.140	1.120 *	1.171	89.07	89.35	90.71	91.06 *	90.93
Circ. 12	1.296	1.291	1.302	1.317	1.288 *	0.729	0.728	0.729	0.754	0.694 *	94.23	94.27	94.17	94.04	94.30 *
Circ. 13	3.192	3.116	3.007	3.021	3.003 *	1.353	1.437	1.241	1.313	1.238 *	91.14	91.55	92.13	92.06	92.15 *
Circ. 14	0.670	0.656	0.595	0.606	0.577 *	0.269	0.275	0.254	0.274	0.243 *	95.23	95.43	96.23	96.10	96.47 *
Circ. 16	4.825	4.461	4.011	3.875 *	3.978	2.912	2.240	2.161	2.154	2.202 *	75.82	79.33	83.29	84.40 *	83.56

Values in bold with an asterisk represent the best results.

Comparatively, based on the results presented in Table 10, the LSTM recurrent neural network model demonstrated superior performance compared to the other models for the majority of the datasets. The LSTM showed good R² values, indicating that it can better estimate the variability in demand patterns compared to the other models. Thus, we assert that the ability of recurrent neural networks to handle temporal and sequential dependencies was beneficial for the task of demand forecasting in the selected circuit datasets. We emphasize that the optimization process conducted to select the best parameters for this model, which are presented in Table 9, was crucial for the achieved performance. On the other hand, the LR model performed the worst among the learning models. This can be attributed to the simplicity of the linear model, which, in most cases, failed to capture complex relationships in the demand data of the selected circuits. In all cases, the RMSE performance followed the results of the R² metric. However, the MAE metric did not always correlate with RMSE and R², as other models generated better results than the LSTM in this evaluation metric.

Regarding the performance of the SVR, RFR, and XGBR models, we can observe in Table 10 that they outperformed the baseline metrics of the LR model. Only in one case, the dataset of circuit 12, did the LR model perform better than the RFR and XGBR models in terms of RMSE, MAE, and R². Depending on the dataset and the selected parameters, at least one of the machine learning models outperformed the others. For circuits 8 and 12, the SVR model stood out among the three models. In circuits 13 and 14, the RFR model performed better than the other two models. For circuits 0, 6, 10, and 16, the XGBR, being more complex than SVR and RFR, achieved better performance. For the datasets of circuits 10 and 16, the XGBR outperformed the LSTM model, which performed better than all the other models for the other datasets. In general, we can observe that the RFR and XGBR models tend to have better performance when compared to SVR in terms of RMSE and MAE in most cases, with XGBR standing out.

Considering the descriptive statistical data presented in Table 7 and Figure 9, we can observe that the variability in average values, standard deviation, and data range of demand influences the performance of the models. In the datasets of circuits 12 and 13, for example, where there is a greater variation in the data range, the SVR and RFR models

outperformed others due to their better handling of data dispersions in these datasets. For the circuit 8 data, where abnormalities (outliers) are illustrated in Figure 9, it was observed, through the R^2 metric in Table 10, that the learning models' generalization ability was significantly affected for this dataset. Additionally, in the circuit 0 dataset, which exhibits greater variations as it represents the entire installation's energy demand, we observed the highest error values. This observation also justifies the performance of the LR models, which are sensitive to outliers, variance, and complex relationships within the datasets. In such cases, more complex and flexible models, such as LSTM, might be needed for capturing demand patterns. It is important to highlight that, to enhance the performance of the LSTM networks considering the high variance of the datasets exposed in Figure 9, we observed that the Optuna optimizer sought to increase the number of LSTM units, as presented in Table 9, so that the learning model could better capture the demand patterns.

Additionally, Table 11 presents the total optimization time for each model to search for the best parameters with the Optuna framework. Subsequently, using the optimal parameters, Table 12 illustrates the training and prediction times for each learning model.

Table 11. Total study time to optimize learning models.

Demand Dataset	Total Study Time (s)			
	SVR	RFR	XGBR	LSTM
Circ. 0	507.04	515.30	1512.57	55,554.55
Circ. 6	488.40	556.79	999.02	22,786.94
Circ. 8	505.96	563.52	1308.05	3995.51
Circ. 10	537.58	605.69	1113.27	24,081.63
Circ. 12	499.79	501.52	1231.30	20,654.27
Circ. 13	521.80	508.32	891.77	27,640.51
Circ. 14	488.50	453.44	1229.87	27,559.57
Circ. 16	555.51	562.22	1333.36	18,281.69

Table 12. Training time and prediction time of learning models.

Demand Dataset	Training Time (ms)					Prediction Time (ms)				
	LR	SVR	RFR	XGBR	LSTM	LR	SVR	RFR	XGBR	LSTM
Circ. 0	35.06	228.90	117.07	3302.61	29,469.54	3.02	3.00	1.94	2.00	281.49
Circ. 6	2.00	126.98	209.68	1740.05	63,918.14	0.99	1.00	3.99	1.99	297.31
Circ. 8	1.01	106.73	31.67	2362.45	64,716.04	1.04	2.00	0.99	1.99	280.51
Circ. 10	2.10	218.34	245.26	1524.37	52,318.09	1.06	2.01	4.84	1.99	668.13
Circ. 12	0.99	61.58	53.69	1804.42	67,387.26	1.14	2.98	2.01	1.00	293.59
Circ. 13	1.99	193.57	105.97	965.42	23,293.50	1.01	3.00	1.05	1.00	269.96
Circ. 14	1.94	184.85	22.01	1186.26	44,027.02	0.99	1.00	0.99	1.51	272.30
Circ. 16	0.99	301.30	192.10	2015.23	24,600.22	1.01	2.01	5.01	2.00	295.66

Despite delivering the highest performance, the LSTM recurrent network model demanded a greater computational time for optimization, training, and prediction processes. As outlined previously in Section 6.7, the variables such as units, batch size, and learning rate significantly influenced the training duration of the LSTM models. On the other hand, the LR model demonstrated a shorter training and prediction timeframe. It is worth noting that the optimization, training, and prediction durations directly correlate with the parameters employed in the model implementation, which varied throughout the hyperparameter tuning process and the learning models' evaluation. For instance, the training time for the RFR model increased for datasets where the tree count was higher, similar to the XGBR model when comparing the results in Table 12 with the hyperparameters displayed in Table 8. In the case of SVR, the regularization parameter C directly impacted the training duration. The XGBR model occupied the second-longest computational time in the training process, while the SVR and RFR models alternated

between the measured durations during the analysis. Hence, for demand data where the training parameters demanded a larger computational effort, the models' training time was extended, subsequently influencing the optimization time for the selected dataset. It is crucial to underscore that, as per Section 6.4.4, although the XGBR model necessitated more training time, its prediction duration was reduced, aligning it closely with simpler models such as LR.

7.2. Evaluation of Our Proposal for Demand Forecast

Table 13 outlines the count of actual demand exceedances beyond 120 kW sourced from the building installation's test data subset (circuit 0), alongside the number of demand exceedances forecasted by each learning model throughout the period from 25 March to 12 April 2023, representing the test set of demand data.

Table 13. Actual and predicted number of demand overruns by learning models.

Actual	LR	SVR	RFR	XGBR	LSTM
38	22	24	30	30	32

As demonstrated, during the testing period for the implemented models, the LSTM model, notwithstanding its higher computational cost for training, proved more effective than other models in the forecasting task. This makes it ideal for use in the SCC to predict the energy demand for the upcoming 15-min intervals in order to avoid demand exceedances. In this context, the LR and SVR models fell short in detecting these exceedances, while the RFR and XGBR models exhibited similar performance. Consequently, the metrics and results elaborated in the prior section align with the comparison made in Table 13.

For comparison purposes, Figure 13 depicts the predictions made by the examined models from 1:00 a.m. on 7 April to 8:00 a.m. on 8 April 2023. The figure highlights the precision with which the models forecast the demand, particularly during periods of minimal variation. Generally speaking, it is observed that the LR, RFR, and SVR models tend to be less precise during moments of variation in comparison to the XGBR and LSTM models. However, during instances of high variation, such as shown for the data from circuit 16, the models are prone to consistent errors that impair their performance in achieving forecasting metrics. Additionally, Figure 14 showcases both actual and forecasted demands using the LSTM neural network models for each circuit's test sets during the period from 26 March to 4 April 2023. For the data from circuits 10, 13, and 16, we highlighted periods of high variance in energy demand in yellow, where the LSTM model did not perform adequately. This situation might be prevalent for loads with constant energy demand variation, as in the case of the three air conditioning units in the installation. Under these circumstances, the RMSE metric penalizes the performance of learning models sensitive to these variations. Consequently, a similar outcome is reflected in the R^2 metric since the model fails to accurately capture these variations. To mitigate these inaccuracies, we could contemplate incorporating other correlated data or different forecasting techniques to enhance the predictability of the forecasting models.

For the circuit 0 data, which represents the entire building installation, we marked in dashed red lines the contracted demand of 120 kW, as shown in Figure 14. From April 1, we observed that the installation's demand exceeded the contracted demand in certain periods. These demand exceedance events are marked in dark red in the figure, both for the installation data (circuit 0) and for the data from the other circuits. We also highlighted in light red the periods in which the circuits had increased demand compared to the data observed in previous periods. We noticed that the algorithm generated forecasts that closely tracked the actual values over time. We suggest using these forecasts to guide the control of the installation's demand and avoid potential exceedances.

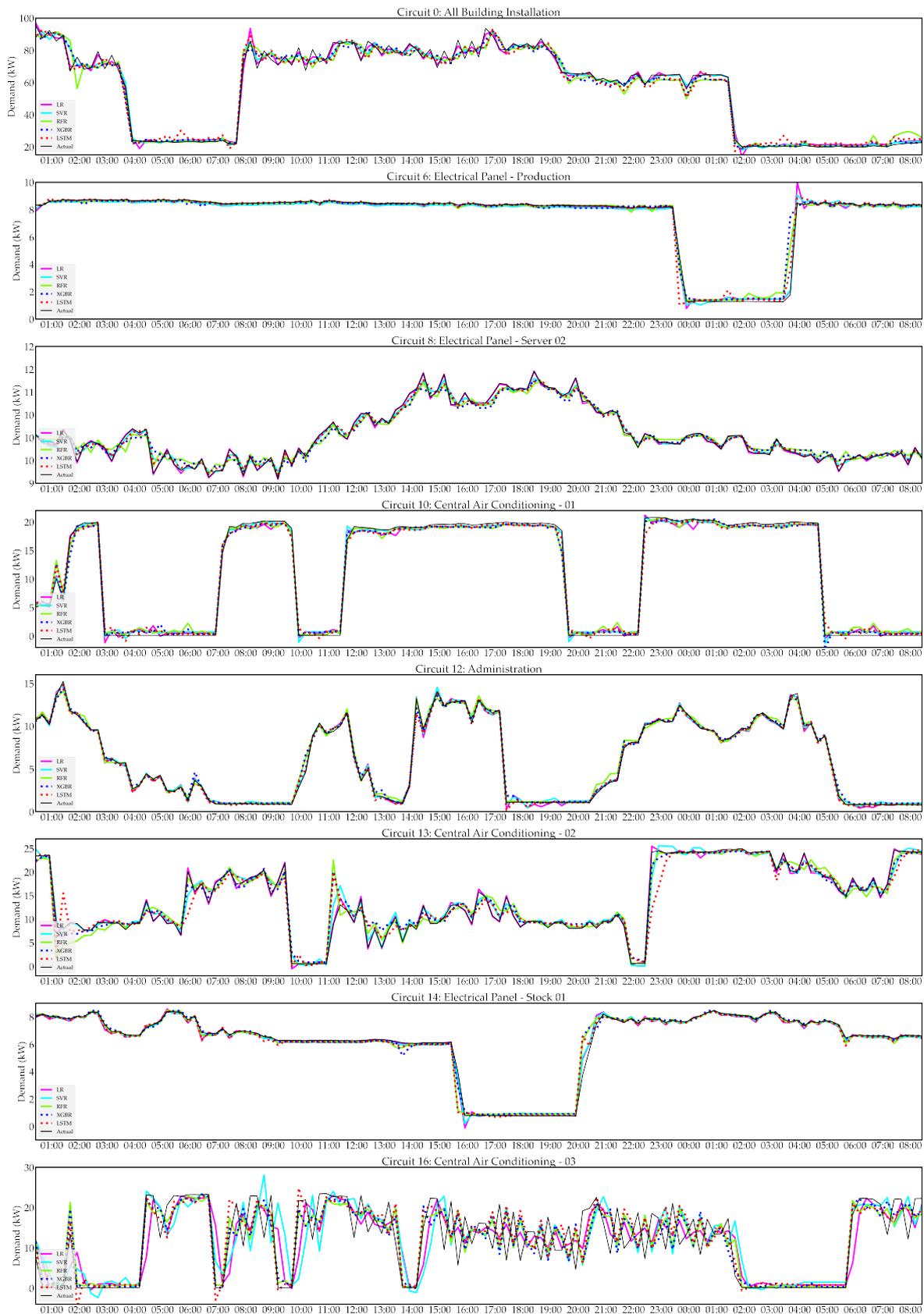


Figure 13. Actual (black) and predicted demand by the LR (magenta), SVR (cyan), RFR (green), XGBR (blue dashed), and LSTM (red dashed) models during the period 01:00 a.m. on 7 April 2023 until 08:00 a.m. of 8 April 2023.

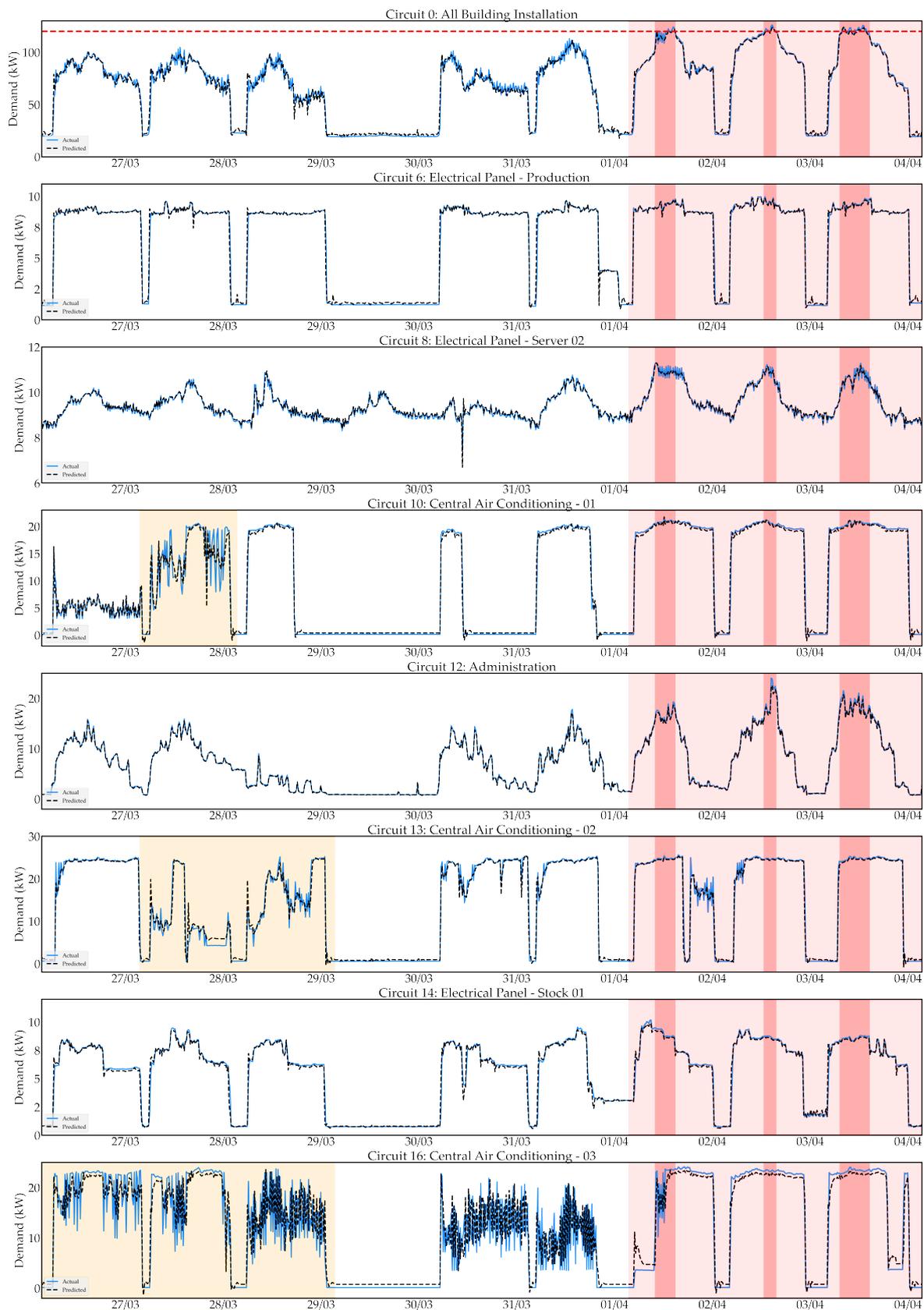


Figure 14. Contracted demand (dashed red), and actual (blue) and predicted values (dashed black) for the 15-min power demand of the selected circuits using the respective proposed LSTM recurrent network models in the period from 26 May to 4 April 2023.

7.3. Discussion of the Results Obtained from the Monitoring Proposal

We implemented a cluster of sensor devices that communicate within a power distribution panel using an ad hoc wireless network. These devices transmit electrical parameters from a building installation and its circuits to a local server, and subsequently to a supervision and control center (SCC). Our proposal's development was based on SmartLVGrid metamodel, which advocates technological updates through the retrofitting of existing systems. To implement the middleware layer of this model, we designed two energy monitoring devices: the ACU-MAIN and the ACU-BREAKER. The ACU-MAIN is responsible for monitoring the main power bus of the installation's distribution panel and acts as a concentrator for the ACU-BREAKER cluster, which monitors the energy consumption of the remaining circuits in the panel.

During the implementation of the ACU-BREAKER and ACU-MAIN devices, we took into account the physical space constraints available in the panel for installation. Therefore, we proposed a novel approach for retrofitting breakers by updating the ACU-BREAKER device compared to the work presented in [5]. This approach facilitates the physical connection interface with the monitoring device, enabling the digital convergence of legacy infrastructure to the smart buildings paradigm. Additionally, we implemented an interoperability layer using request and response message exchanges that travel through the physical layer of the IEEE 802.11 standard via the ESP-NOW protocol. This wireless communication enables our retrofitting proposal without the need for additional wired ethernet network points, following the directives of the factory in which our study took place. Thus, we enable flexible retrofitting of the installation by leveraging pre-existing resources and adding capabilities to enable energy management.

Our proposal has been operating continuously and uninterruptedly since the start of data collection after its installation, validating our approach to building energy monitoring retrofitting. As a result, we were able to build a database containing energy data from the legacy installation for its managers, including power factor, active energy, current, and voltage data for both the overall installation and individual circuits. This has enabled data-driven energy management of the legacy installation, as the monitored data became available in databases and dashboards at the supervision and control center (SCC).

7.4. Discussion of the Results Obtained for Forecasting Energy Demand in the Proposed Scenario

Based on the Brazilian regulatory resolution ANEEL n° 1000/2021 [57], the consumer unit in question falls under the binomial tariff structure. In this case, it is charged based on both consumption and a contracted limit demand, which is measured by the energy utility every 15 min. Incidentally, during periods of high production, the factory exceeds the contracted demand of 120 kW and consequently incurs penalties. With the collected database, we conducted an analysis of the loads that contribute the most to the increase in consumption and demand exceedances of the installation using Pareto analysis. We identified seven loads that contribute to nearly 80% of the total installation consumption. Based on this, we analyzed the variations in energy demand every 15 min for the loads of these circuits. To perform our analysis, we applied the sliding window technique with 10 previous demand samples and min–max normalization as a processing step for demand forecasting for the next 15 min. Subsequently, we employed various learning models, namely, linear regression (LR), support vector regressor (SVR), random forest regressor (RFR), XGBoost regressor (XGBR), and a long short-term memory (LSTM) recurrent neural network model. We evaluated the performance of each model and, to ensure the best possible performance, we utilized the optimization framework Optuna to search for the best parameters for the demand data of each selected circuit.

We observed that the LSTM model performed the best, followed by the XGBR, RFR, and SVR models, respectively. The LSTM model was able to capture the demand pattern of the selected circuits most effectively, as shown in the metrics presented in Table 10, and it predicted the highest number of demand exceedances for the test set, as shown in Table 13. However, the LSTM model required the longest computation time for optimization, training,

and making predictions (Tables 11 and 12). All the other models outperformed the baseline LR metrics, with notable performance from the XGBR model, which outperformed LSTM for two datasets (circuits 10 and 16). This opens up opportunities for future neural network architectures that can surpass the metrics presented in Table 10. In Figure 14, we can observe that the predictions made by the LSTM model performed well for the selected circuit datasets. We noted that depending on the nature of the monitored loads, there may be data variations that could affect the predictability of the forecasting algorithms. We hope that by increasing the dataset size and incorporating other variables correlated with demand and seasonality, we can improve the performance of the learning algorithms for demand forecasting tasks. In our research, we have achieved the objective of demonstrating the impact and relevance of monitoring and forecasting the energy demand of circuits in a legacy building installation, aiming to detect possible breaches of contracted demand and identify the circuits where action should be taken to rectify demand transgressions in line with the regulatory framework of the Brazilian energy system.

8. Conclusions

In this work, we developed an AIoT strategy that performs energy demand forecasting for a legacy building installation and its circuits for the next 15 min, based on the retrofit of the pre-existing energy system and the premises of the SmartLVGrid metamodel. The protocols of the SmartLVGrid metamodel enabled us to design an architecture that facilitates the technological transformation of a legacy installation into the smart buildings paradigm, making the most of the existing resources.

During the development of this study, we conceived a cluster of sensor devices called ACU-BREAKERS that monitor the individual electrical parameters of each electrical circuit and communicate through an ESP-NOW ad hoc network with a coordinating device called ACU-MAIN. In our proposal, the ACU-MAIN device performs multiple functions, including coordinating data requests from other ACUs, monitoring the main power bus of the installation, and transmitting the collected data via ethernet to a locally available server within the installation. The server, in turn, forwards the collected data to the cloud-hosted SCC, where data analysis is conducted to improve the energy management processes.

Our proposal operated continuously from 15 January to 12 April 2023, and with the data obtained we conducted statistical analyses to identify the loads that contributed the most to the increase in consumption and energy demand of the installation. Based on Brazilian regulations, we focused on forecasting for the next 15 min to detect possible demand surpluses in the installation and identify the main loads causing this transgression. In this way, we provided data-driven insights for decision making regarding possible surpluses and where and when to act to control the load demand.

We employed preprocessing techniques such as sliding window for dividing the training and testing datasets of each circuit, along with min–max normalization of the data. As learning models, we used LR as the baseline for evaluating the machine learning models SVR, RFR, XGBR, and an LSTM-based recurrent neural network model. The hyperparameters of each learning model were optimized using the Optuna framework for Bayesian optimization, in order to extract the best possible performance. Subsequently, we evaluated the learning models, and the LSTM model outperformed the other learning models, followed by XGBR, RFR, SVR, and LR. In this order, the models had longer training and optimization times. We also evaluated which models successfully predicted the highest number of demand surpluses, with a highlight on the LSTM and XGBR models.

It is important to emphasize that we evaluated a model for each dataset of each circuit. For the construction of building electrical systems with more circuits and power boards, the implementation of learning models for each dataset could become unfeasible. In addition, for other cases and systems, the use of other learning models, preprocessing, and feature selection methods and other retrofit strategies could be adopted to obtain better results for the benefit of a more sustainable building ecosystem.

However, whether to optimize the use of energy inputs or to plan operations in building facilities, in our proposal, the forecast and monitoring of energy demand allow data-based management of pre-existing energy systems in legacy facilities. In precarious scenarios, without infrastructure or resources to implement modern control and communication systems, our retrofit architecture facilitates a non-abrupt digital transformation towards smart building convergence, leveraging AIoT concepts and predictive models based on wireless network data. In addition, we digitized the installation's circuits using the assumptions of our retrofit architecture, which recommends taking advantage of existing resources through well-defined protocol stacks. We emphasize that the proposed architecture represents an alternative for using electrical parameters from legacy circuits to create databases for predictive analysis, such as the energy demand forecast presented in this work. Thus, it is possible to guarantee the sustainability and improve the energy efficiency of old building installations.

9. Future Perspectives

Once we make the electrical system observable and allocate resources for demand forecasting, we enable the management of current and future energy resources from the demand side. Therefore, for future work, we suggest allocating local intelligence resources to implement new strategies that include demand control of the installation based on local business rules. This can be achieved by controlling the loads present in the installation's circuits, as we know which loads will affect the installation during demand exceedances. By also forecasting the demand of the installation's loads, we suggest utilizing distributed energy resources to inject the necessary energy to compensate for the energy demand during peak moments, avoiding possible exceedances from the energy generation side. In this way, renewable or non-renewable resources can be activated based on the proposed predictive intelligence to partially or fully meet the installation's energy demand.

Additionally, we suggest that this process may involve new dynamic energy markets, where energy sources from free energy markets can be negotiated and utilized depending on the predictability scenario of demand exceedances to reduce the costs associated with possible exceedances. The prediction task can also analyze future energy costs, recommending potential energy suppliers based on this dynamic analysis. Further work in this field can explore other prediction resources based on other energy aspects of a building installation, involving protection systems, energy consumption, or power quality. This includes studies focused on optimizing energy utilization and mitigating harmonics in the installation.

From the perspective of artificial intelligence models, we suggest evaluating the proposed strategy for other learning model architectures and datasets, including variations of the LSTM recurrent neural network model in the context of building electrical circuits in smart buildings. We also recommend using other preprocessing techniques and different sliding window sizes to assess the performance of the learning models in short, medium, and long-term prediction contexts, depending on the study's needs. For future work, we suggest exploring knowledge transfer techniques to facilitate the training of other learning models for circuits within the same cluster and for clusters located in other locations or installations. In this work, we developed specialized demand forecasting models for each circuit of the installation, which can make it costly to maintain the system in some cases. Through knowledge transfer techniques, it is possible to generalize the demand pattern capturing techniques for circuits in a building installation and scale this strategy to other cases and systems, involving the same installation or other legacy installations.

Author Contributions: Conceptualization, R.A.F.; methodology, R.A.F., R.C.S.G. and C.T.C.J.; software, R.A.F.; validation, R.A.F.; formal analysis, R.A.F., C.C., N.L.V., L.B.F.N., I.G.T. and F.R.S.; investigation, R.A.F.; resources, R.A.F.; data curation, R.A.F.; writing—original draft preparation, R.A.F.; writing—review and editing, C.C., N.L.V., L.B.F.N., H.L.N.d.S. and F.R.S.; visualization, R.A.F.; supervision, R.C.S.G. and C.T.C.J.; project administration, R.C.S.G., H.L.N.d.S. and I.G.T.; funding acquisition, I.G.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: The authors thank the Embedded Systems Laboratory of the State University of Amazonas for the support provided.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACU	Automation and communication unit
AEMO	Australian Energy Market Operator
AI	Artificial intelligence
AIoT	Artificial intelligence of things
ANEEL	Agência Nacional de Energia Elétrica
ANN	Artificial neural network
AR	Autoregressive model
ARIMA	Autoregressive integrated moving average
API	Application programming interface
BGM	Bayesian Gaussian mixture
CIN	Coupling and interaction node
Cire.pl	Centrum Informacji o Rynku Energii
CNN	Convolutional neural networks
CSFs	Computational support functions
Damas	Damas Energy information system
DLM	Dynamic linear model
DRFs	Domain retrofitting functions
DTR	Decision tree regression
ECMWF	European Centre for Medium-Range Weather Forecasts
EIA	Energy Information Administration
EM-GMM	Expectation maximization Gaussian mixture model
ENR	Elastic net regression
ETS	Smoothing state space model
ES	Exponential smoothing
FAR	Functional autoregressive model
FARX	Fractional-order autoregressive model with exogenous variables
FCC	Florida Climate Center
FFANN	Feedforward artificial neural network
GBR	Gradient boosting regression
GPU	Graphics processing unit
GRNN	General regression neural network
GRU	Gated recurrent unit
HW	Holt–Winters
IEEE	Institute of Electrical and Electronics Engineers
IESO	Independent electricity system operator
IoT	Internet of things
ISFs	Interdomain support functions
JSON	JavaScript object notation
KEPCO	Korea Electric Power Corporation
KMA	Korea Meteorological Administration
KNNR	K-nearest neighbor regression

LAN	Local area network
LR	Linear regression
LSTM	Long short-term memory
MAC	Media access control
MAE	Mean absolute error
MAN	Metropolitan area network
MLP	Multilayer perceptron
MPR	Multivariate polynomial regression
MRM	Multiple regression model
MQTT	Message queue telemetry transport
NARX	Non-linear autoregressive exogenous
N-BEATS	Neural basis expansion analysis for interpretable time series
NNAR	Autoregressive neural networks
NNETAR	Neural network time series forecasts
ONS	Operador Nacional do Sistema
OPs	Operational primitives
OPSD	Open power system data
P2P	Peer-to-peer
PoI	Points of interface
PR	Polynomial regression
QoS	Quality of service
R ²	R-squared score
RAM	Random access memory
RFR	Random forest regressor
RMS	Root mean square
RMSE	Root mean squared error
RNN	Recurrent neural networks
RS	Regression with seasonality
SARIMA	Seasonal ARIMA
SCC	Supervision and control center
SLFN	Single-layer feedforward neural networks
SmartLVGrid	Smart low-voltage grids
SN	Service node
SoC	System-on-a-chip
SSD	Solid state drive
SVR	Support vector regression
TBATS	Trigonometric Box–Cox transform, ARMA errors, trend, and seasonal components
TCN	Temporal convolutional network
TCP	Transmission control protocol
TFT	Temporal fusion transformer
U.S.	United States
W	Watts
WSN	Wireless sensor network
XGBoost	Extreme gradient boosting
XGBR	XGBoost regressor

References

1. Wen, Y.; Fashiar Rahman, M.; Xu, H.; Tseng, T.L.B. Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective. *Measurement* **2022**, *187*, 110276. [[CrossRef](#)]
2. Chatterjee, S.; Chaudhuri, R.; Shah, M.; Maheshwari, P. Big data driven innovation for sustaining SME supply chain operation in post COVID-19 scenario: Moderating role of SME technology leadership. *Comput. Ind. Eng.* **2022**, *168*, 108058. [[CrossRef](#)] [[PubMed](#)]
3. Gomes, R.C.S.; Costa, C.; Silva, J.; Sicchar, J. SmartLVGrid Platform—Convergence of Legacy Low-Voltage Circuits toward the Smart Grid Paradigm. *Energies* **2019**, *12*, 2590. [[CrossRef](#)]
4. Fernandes, R.A.; Gomes, R.C.S.; Dias, O.; Carvalho, C. A Novel Strategy for Smart Building Convergence Based on the SmartLVGrid Metamodel. *Energies* **2022**, *15*, 1016. [[CrossRef](#)]

5. Fernandes, R.A.; Gomes, R.C.S.; Dias, O.; Carvalho, C.; Torné, I.G.; Oliveira, J.P.; Júnior, C.T.C. A Retrofit Strategy for Real-Time Monitoring of Building Electrical Circuits Based on the SmartLVGrid Metamodel. *Energies* **2022**, *15*, 9234. [[CrossRef](#)]
6. Yu, Z.; Khan, S.A.R.; Ponce, P.; ul haq, H.M.Z.; Ponce, K. Exploring essential factors to improve waste-to-resource recovery: A roadmap towards sustainability. *J. Clean. Prod.* **2022**, *350*, 131305. [[CrossRef](#)]
7. Amin, N.; Song, H.; Khan, Z.A. Dynamic linkages of financial inclusion, modernization, and environmental sustainability in South Asia: A panel data analysis. *Environ. Sci. Pollut. Res.* **2022**, *29*, 16588–16596. [[CrossRef](#)]
8. Bronner, W.; Gebauer, H.; Lamprecht, C.; Wortmann, F. Sustainable AIoT: How artificial intelligence and the internet of things affect profit, people, and planet. In *Connected Business: Create Value in a Networked Economy*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 137–154.
9. El Himer, S.; Ouaisa, M.; Ouaisa, M.; Boulouard, Z. Artificial Intelligence of Things (AIoT) for Renewable Energies Systems. In *Artificial Intelligence of Things for Smart Green Energy Management*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–13.
10. da Silva Mendonça, R.; de Oliveira Lins, S.; de Bessa, I.V.; de Carvalho Ayres, F.A.; de Medeiros, R.L.P.; de Lucena, V.F. Digital Twin Applications: A Survey of Recent Advances and Challenges. *Processes* **2022**, *10*, 744. [[CrossRef](#)]
11. Kröll, M.; Cseh, C. Implementation Model for Digital Retrofit for Sustainable Production. *Procedia Comput. Sci.* **2023**, *217*, 486–494. [[CrossRef](#)]
12. Silva, D.S.; Nascimento, L.B.F.; Fernandes, R.A.; Gomes, R.C.S.; Torné, I.G. Arquitetura para identificar e estimar regiões de faltas permanentes em média tensão: Uma Contribuição da Plataforma SmatLVGrid/Architecture to identify and estimate regions of permanent faults in medium voltage: A Contribution of the SmatLVGrid Platform. *Braz. J. Dev.* **2021**, *7*, 24845–24860. [[CrossRef](#)]
13. Zielińska-Sitkiewicz, M.; Chrzanowska, M.; Furmańczyk, K.; Pacutkowski, K. Analysis of Electricity Consumption in Poland Using Prediction Models and Neural Networks. *Energies* **2021**, *14*, 6619. [[CrossRef](#)]
14. Velasquez, C.E.; Zocатели, M.; Estanislau, F.B.; Castro, V.F. Analysis of time series models for Brazilian electricity demand forecasting. *Energy* **2022**, *247*, 123483. [[CrossRef](#)]
15. Leite Coelho da Silva, F.; da Costa, K.; Canas Rodrigues, P.; Salas, R.; López-Gonzales, J.L. Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector. *Energies* **2022**, *15*, 588. [[CrossRef](#)]
16. Shah, I.; Jan, F.; Ali, S. Functional data approach for short-term electricity demand forecasting. *Math. Probl. Eng.* **2022**, *2022*, 6709779. [[CrossRef](#)]
17. Manno, A.; Martelli, E.; Amaldi, E. A Shallow Neural Network Approach for the Short-Term Forecast of Hourly Energy Consumption. *Energies* **2022**, *15*, 958. [[CrossRef](#)]
18. Rajula, H.S.R.; Verlató, G.; Manchia, M.; Antonucci, N.; Fanos, V. Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment. *Medicina* **2020**, *56*, 455. [[CrossRef](#)]
19. Pavlicko, M.; Vojteková, M.; Blažeková, O. Forecasting of Electrical Energy Consumption in Slovakia. *Mathematics* **2022**, *10*, 577. [[CrossRef](#)]
20. Aisyah, S.; Simaremare, A.A.; Adytia, D.; Aditya, I.A.; Alamsyah, A. Exploratory Weather Data Analysis for Electricity Load Forecasting Using SVM and GRNN, Case Study in Bali, Indonesia. *Energies* **2022**, *15*, 3566. [[CrossRef](#)]
21. Shirzadi, N.; Nizami, A.; Khazen, M.; Nik-Bakht, M. Medium-Term Regional Electricity Load Forecasting through Machine Learning and Deep Learning. *Designs* **2021**, *5*, 27. [[CrossRef](#)]
22. Arjomandi-Nezhad, A.; Ahmadi, A.; Taheri, S.; Fotuhi-Firuzabad, M.; Moeini-Aghtaie, M.; Lehtonen, M. Pandemic-Aware Day-Ahead Demand Forecasting Using Ensemble Learning. *IEEE Access* **2022**, *10*, 7098–7106. [[CrossRef](#)]
23. Rawal, K.; Ahmad, A. A Comparative Analysis of Supervised Machine Learning Algorithms for Electricity Demand Forecasting. In Proceedings of the 2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T), Raipur, India, 1–3 March 2022; pp. 1–6. [[CrossRef](#)]
24. Wang, Y.; Fu, Z.; Wang, F.; Li, K.; Li, Z.; Zhen, Z.; Dehghanian, P.; Fotuhi-Firuzabad, M.; Catalão, J.P.S. Adaptive Optimal Greedy Clustering-Based Monthly Electricity Consumption Forecasting Method. *IEEE Trans. Ind. Appl.* **2022**, *58*, 7881–7891. [[CrossRef](#)]
25. Farrokhbadi, M.; Browell, J.; Wang, Y.; Makonin, S.; Su, W.; Zareipour, H. Day-Ahead Electricity Demand Forecasting Competition: Post-COVID Paradigm. *IEEE Open Access J. Power Energy* **2022**, *9*, 185–191. [[CrossRef](#)]
26. Bashir, T.; Haoyong, C.; Tahir, M.F.; Liqiang, Z. Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN. *Energy Rep.* **2022**, *8*, 1678–1686. [[CrossRef](#)]
27. Elkamel, M.; Schleider, L.; Pasiliao, E.L.; Diabat, A.; Zheng, Q.P. Long-Term Electricity Demand Prediction via Socioeconomic Factors—A Machine Learning Approach with Florida as a Case Study. *Energies* **2020**, *13*, 3996. [[CrossRef](#)]
28. Torres, J.; Martínez-Álvarez, F.; Troncoso, A. A deep LSTM network for the Spanish electricity consumption forecasting. *Neural Comput. Appl.* **2022**, *34*, 10533–10545. [[CrossRef](#)]
29. Mustaqeem; Ishaq, M.; Kwon, S. Short-Term Energy Forecasting Framework Using an Ensemble Deep Learning Approach. *IEEE Access* **2021**, *9*, 94262–94271. [[CrossRef](#)]
30. Nazir, A.; Shaikh, A.K.; Shah, A.S.; Khalil, A. Forecasting energy consumption demand of customers in smart grid using Temporal Fusion Transformer (TFT). *Results Eng.* **2023**, *17*, 100888. [[CrossRef](#)]
31. Shaikh, A.K.; Nazir, A.; Khan, I.; Shah, A.S. Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series. *Sci. Rep.* **2022**, *12*, 22562. [[CrossRef](#)]
32. Nabavi, S.A.; Motlagh, N.H.; Zaidan, M.A.; Aslani, A.; Zakeri, B. Deep Learning in Energy Modeling: Application in Smart Buildings With Distributed Energy Generation. *IEEE Access* **2021**, *9*, 125439–125461. [[CrossRef](#)]

33. Li, W. Application of Economical Building Management System for Singapore Commercial Building. *IEEE Trans. Ind. Electron.* **2020**, *67*, 4235–4243. [[CrossRef](#)]
34. Eseye, A.T.; Lehtonen, M.; Tukia, T.; Uimonen, S.; John Millar, R. Machine Learning Based Integrated Feature Selection Approach for Improved Electricity Demand Forecasting in Decentralized Energy Systems. *IEEE Access* **2019**, *7*, 91463–91475. [[CrossRef](#)]
35. Lee, H.; Kim, D.; Gu, J.H. Prediction of Food Factory Energy Consumption Using MLP and SVR Algorithms. *Energies* **2023**, *16*, 1550. [[CrossRef](#)]
36. Mounter, W.; Ogwumike, C.; Dawood, H.; Dawood, N. Machine Learning and Data Segmentation for Building Energy Use Prediction—A Comparative Study. *Energies* **2021**, *14*, 5947. [[CrossRef](#)]
37. Durand, D.; Aguilar, J.; R-Moreno, M.D. An Analysis of the Energy Consumption Forecasting Problem in Smart Buildings Using LSTM. *Sustainability* **2022**, *14*, 13358. [[CrossRef](#)]
38. Mariano-Hernández, D.; Hernández-Callejo, L.; Solís, M.; Zorita-Lamadrid, A.; Duque-Pérez, O.; Gonzalez-Morales, L.; García, F.S.; Jaramillo-Duque, A.; Ospino-Castro, A.; Alonso-Gómez, V.; et al. Analysis of the Integration of Drift Detection Methods in Learning Algorithms for Electrical Consumption Forecasting in Smart Buildings. *Sustainability* **2022**, *14*, 5857. [[CrossRef](#)]
39. Arivukkody, V.; Gokulakannan, T.; Kalpana, S. Aiot Based Residential Smart Energy Meter with Power Saving Methodology. In Proceedings of the 2022 1st International Conference on Computational Science and Technology (ICCST), Chennai, India, 9–10 November 2022; pp. 80–85. [[CrossRef](#)]
40. Chandra Das, N.; Ziaul Haque Zim, M.; Sazzad Sarkar, M. Electric Energy Meter System Integrated with Machine Learning and Conducted by Artificial Intelligence of Things—AioT. In Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), St. Petersburg, Russia, 26–28 January 2021; pp. 826–832. [[CrossRef](#)]
41. Salama, A.K.; Abdellatif, M.M. AIoT-based Smart Home Energy Management System. In Proceedings of the 2022 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT), Alamein New City, Egypt, 18–21 December 2022; pp. 177–181. [[CrossRef](#)]
42. Kumar, L.; Choudhury, D.; Paduri, A.R.; Kumar, S.; Sahoo, D.; Murthy, J.; Darapaneni, N. Electric Vehicle (EV) Preventive Diagnostic System: Solution for Thermal Management of Battery packs using AIOT. In Proceedings of the 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–11 March 2023; pp. 0041–0046. [[CrossRef](#)]
43. Khanchuea, K.; Siripokarpirom, R. A Multi-Protocol IoT Gateway and WiFi/BLE Sensor Nodes for Smart Home and Building Automation: Design and Implementation. In Proceedings of the 2019 10th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES), Bangkok, Thailand, 25–27 March 2019; pp. 1–6. [[CrossRef](#)]
44. Abdul, M.S.; Sam, S.M.; Mohamed, N.; Hassan, N.H.; Azizan, A.; Yusof, Y.M. Peer to Peer Communication for the Internet of Things Using ESP32 Microcontroller for Indoor Environments. In Proceedings of the 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Republic of Korea, 19–21 October 2022; pp. 1–6. [[CrossRef](#)]
45. Eridani, D.; Rochim, A.F.; Cesara, F.N. Comparative Performance Study of ESP-NOW, Wi-Fi, Bluetooth Protocols based on Range, Transmission Speed, Latency, Energy Usage and Barrier Resistance. In Proceedings of the 2021 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 18–19 September 2021; pp. 322–328. [[CrossRef](#)]
46. Hoang, T.N.; Van, S.T.; Nguyen, B.D. ESP-NOW Based Decentralized Low Cost Voice Communication Systems For Buildings. In Proceedings of the 2019 International Symposium on Electrical and Electronics Engineering (ISEE), Ho Chi Minh City, Vietnam, 10–12 October 2019; pp. 108–112. [[CrossRef](#)]
47. Espressif. ESP32 Series Datasheet. 2022. Available online: https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf (accessed on 10 July 2023).
48. Espressif. ESP32-WROOM-32E and ESP32-WROOM-32UE Datasheet. 2022. Available online: https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32e_esp32-wroom-32ue_datasheet_en.pdf (accessed on 10 July 2023).
49. Analog Devices. Datasheet ADE7758—Poly Phase Multifunction Energy Metering IC with Per Phase Information. 2011. Available online: <https://www.analog.com/media/cn/technical-documentation/data-sheets/ADE7758.pdf> (accessed on 10 July 2023).
50. Sanjuan, E.B.; Cardiel, I.A.; Cerrada, J.A.; Cerrada, C. Message Queuing Telemetry Transport (MQTT) Security: A Cryptographic Smart Card Approach. *IEEE Access* **2020**, *8*, 115051–115062. [[CrossRef](#)]
51. Toldinas, J.; Lozinskis, B.; Baranauskas, E.; Dobrovolskis, A. MQTT Quality of Service versus Energy Consumption. In Proceedings of the 2019 23rd International Conference Electronics, Palanga, Lithuania, 17–19 June 2019; pp. 1–4. [[CrossRef](#)]
52. Ohno, S.; Terada, K.; Yokotani, T.; Ishibashi, K. Distributed MQTT broker architecture using ring topology and its prototype. *IEICE Commun. Express* **2021**, *10*, 582–586. [[CrossRef](#)]
53. Accuenergy. AcuCT Hinged Series Datasheet. 2021. Available online: <https://www.accuenergy.com/wp-content/uploads/acuct-hinged-series-compact-split-core-current-transformer-datasheet.pdf> (accessed on 10 July 2023).
54. MTE Meter Test Equipment. PPS 400.3: Three-Phase Portable Power Source (12 A or 120 A/300 V). 2004. Available online: <https://www.acitqatar.com/product/pps-400-3/> (accessed on 10 July 2023).
55. Guimarães, A.; Freitas, T.; Griner, H.; De Almeida, T. Smart energy monitoring system with ADE7758 IC. In Proceedings of the 2015 5th International Youth Conference on Energy (IYCE), Pisa, Italy, 27–30 May 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–5.

56. Rodrigues, V.; Moraes, R.; Berejuck, M. A Brazilian Legal and Technical Evaluation about Energy Binomial Tariff. In Proceedings of the 2021 IST-Africa Conference (IST-Africa), Online, 10–14 May 2021; pp. 1–8.
57. National Agency of Electric Energy (ANEEL)—Normative Resolution No. 1000/2021. Available online: <https://www2.aneel.gov.br/cedoc/ren20211000.pdf> (accessed on 29 September 2022).
58. Wu, B.; Cai, W.; Cheng, F.; Chen, H. Simultaneous-fault diagnosis considering time series with a deep learning transformer architecture for air handling units. *Energy Build.* **2022**, *257*, 111608. [[CrossRef](#)]
59. Hodson, T.O. Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev.* **2022**, *15*, 5481–5487. [[CrossRef](#)]
60. Qi, J.; Du, J.; Siniscalchi, S.M.; Ma, X.; Lee, C.H. On mean absolute error for deep neural network based vector-to-vector regression. *IEEE Signal Process. Lett.* **2020**, *27*, 1485–1489. [[CrossRef](#)]
61. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [[CrossRef](#)] [[PubMed](#)]
62. Rathaur, S.; Kamath, N.; Ghanekar, U. Software defect density prediction based on multiple linear regression. In Proceedings of the 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 15–17 July 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 434–439.
63. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2022.
64. Raschka, S.; Mirjalili, V. *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow 2*; Packt Publishing Ltd.: Birmingham, UK, 2019.
65. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [[CrossRef](#)]
66. Aggarwal, C.C. *Neural Networks and Deep Learning*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 10, p. 3.
67. Ali, Y.A.; Awwad, E.M.; Al-Razgan, M.; Maarouf, A. Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity. *Processes* **2023**, *11*, 349. [[CrossRef](#)]
68. Arden, F.; Safitri, C. Hyperparameter Tuning Algorithm Comparison with Machine Learning Algorithms. In Proceedings of the 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 13–14 December 2022; pp. 183–188. [[CrossRef](#)]
69. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, USA, 4–8 August 2019.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.