Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3, 1892-1907
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i3.5716

Hybrid approach based on wavelets, Kalman filters, Arima, and neural networks for mastering electrical energy at the residential level: Forecasting the electrical energy consumption of residences in the city of Lomé

Yéloiti N'poh Babikou[1], Agbassou Guenoukpati[1,2,3], Adekunlé Akim Salami[1,2*], Koffi Mawugno Kodjo[1,2]
[1]Centre d'Excellence Régional pour la Maîtrise de l'Electricité (CERME), University of Lome, Togo; yeloiti@gmail.com (Y.N.B.); guenoukpatib@gmail.com (A.G.); akim_salami@yahoo.fr (A.A.S.).
[2]Department of Electrical Engineering, Ecole Polytechnique de Lomé (EPL), University of Lome, Togo.
[3]Laboratoire de Recherche en Sciences de l'Ingénieur (LARSI), University of Lome, Togo.

Abstract: The energy sector, specifically residential electricity, must be managed to understand the electricity consumption model of each residence at any given time. This study introduces three hybrid forecasting techniques: Wavelet Transform-Kalman filters-ARIMA (WKA), Wavelet Transform-Artificial Neural Networks-Kalman filters (WNNK), and Wavelet Transform-Artificial Neural Networks-ARIMA (WNNA). These hybrid forecasting models were individually applied to each dataset of each residence. The data for this study were collected from 28 different residences over a period of 2 years and 5 months in Lomé, with measurements taken at one-minute intervals. The results, validated using error evaluation criteria such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the correlation coefficient, revealed that 10 of the 28 residences achieved the best forecasting results with the WNNK hybrid model, 10 with the WKA model, and 8 with the WNNA model. This analysis enabled the classification of residences into three groups: Group 1, Group 2, and Group 3, corresponding respectively to the residences achieving the most accurate forecasting results with the hybrid models WNNK, WNNA, and WKA. This work not only enhanced the understanding of electricity consumption habits and provided a method for forecasting future electricity use but also categorized each residence into one of the three groups based on its level of consumption.
Keywords: *Electrical energy, Forecasting, Hybrid model, Residential consumption.*

1. Introduction

Effective management of residential electrical energy remains a challenge for the development of the energy sector, considering the centralized or decentralized integration of external production sources into the overall energy management system. Proper management of residential electrical energy requires a thorough understanding of each residence's real-time energy consumption habits and accurate forecasting for each residence. When implemented, the integration of sometimes intermittent renewable energy resources and their management within the electrical energy system adds a layer of complexity to the system's operation [1]. Therefore, short-term and medium-term residential electrical energy forecasting is crucial. When done correctly, this type of energy forecasting plays a vital role in improving the reliability and efficiency of the electrical system and simplifies the integration of external energy resources while reducing operating costs [2]. Additionally, the availability of a database containing residential electrical energy data is essential for analyzing, understanding, and modeling residential energy consumption [3, 4]. Using artificial intelligence methods to forecast nonlinear and noisy systems yields better results than statistical methods, which are limited to stationary data.

History: Received: 8 January 2025; Revised: 18 March 2025; Accepted: 21 March 2025; Published: 24 March 2025
* Correspondence: akim_salami@yahoo.fr

Residential electrical energy data is very noisy, complex, and non-stationary. Several artificial intelligence methods have been used to predict residential electrical energy [5-7]. The non-stationarity, complexity, and noise present in electrical energy data can be explained by several factors, such as the unpredictable behavior of consumers towards electrical energy, sudden events, load shedding, and meteorological variations. All these factors limit the accuracy of predictions, even for artificial intelligence methods [8, 9]. Several hybrid methods, considering the strengths and weaknesses of the models to be hybridized, have improved the accuracy of residential electrical energy forecasts. To further enhance the performance of electrical energy forecasts, wavelet transform has often been coupled with the ARIMA model and/or neural networks to form a more efficient and robust hybrid model than singular artificial intelligence models [10]. Signal decomposition using the wavelet method allows data to be separated into several components. After obtaining the optimal decomposition level, an algorithm selects certain components from the decomposition to form hybrid models, while the rest is considered noise and eliminated from the model. Wavelet decomposition is a preprocessing step that improves performance for noisy data, such as residential electrical energy, before combining it with ARIMA models, neural networks, or Kalman filters. Hybrid models based on Kalman filters and/or wavelet transform have been used to improve prediction results in the residential energy sector [11-13]. This study aims to investigate residential electrical energy forecasting. To achieve this, several hybrid models were developed based on either wavelet transform, Kalman filters, or both, combined with a statistical model (ARIMA) and/or a machine learning model (artificial neural networks). The three hybrid models, WKA, WNNK, and WNNA, were applied to the collected residential electrical energy data and evaluated to validate the model that provides the optimal forecast for each considered residence. The forecasting performance results were evaluated using criteria such as the correlation coefficient (R2), root mean square error (RMSE), and mean absolute percentage error (MAPE) for the three proposed hybrid models. The rest of this work is structured into several sections: Section 2 covers the literature review, Section 3 discusses the framework of the hybrid approaches WNNK, WNNA, and WKA, Section 4 shows the performance evaluation criteria and Section 5 presents the results and discussions.

2. Literature Review

Several methods in the literature allow for the prediction of residential electrical energy. For instance, in Ramos, et al. [14] the authors compare the performance of different model architectures, such as recurrent neural networks (RNN), long short-term memory (LSTM), gated recurrent units (GRU), and time series transformers (TST). The study also proposes an ensemble method optimized by simulated annealing to improve the accuracy of residential energy consumption forecasts. La Tona, et al. [15] employs an LSTM encoder-decoder architecture to integrate historical and future exogenous variables. The results were compared to three other methods, showing a reduction in mean absolute error of up to 8%. Hybrid methods provide additional improvements in prediction, as seen in Fan, et al. [16] where the authors propose a hybrid method combining empirical wavelet transform, long short-term memory, and support vector machines. The results show that this hybrid model outperforms other models. Gao, et al. [17] proposes an innovative hybrid model to predict residential electricity consumption. This model integrates online search data to improve forecast accuracy. The study results demonstrate a significant improvement in forecast accuracy compared to traditional models. Research related to prediction using hybrid methods formed by combining three of the following models: wavelets, Kalman filters, neural networks, and the ARIMA model, is not related to residential prediction. Patel and Deb [18] developed several hybrid models for wind forecasting based on the Kalman filter and wavelet transform. Their combination with the ARIMA model or certain machine learning models such as Support Vector Regression and Random Forest Regression revealed that the best precise prediction performance was found with the hybrid Kalman filter-wavelet transform-machine learning model (KF-WT-ML) on different terrains. In Khashei and Mahdavi Sharif [19] proposed a hybrid model combining a Kalman filter to preprocess data and reduce noise, ARIMA, and artificial

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

neural networks to improve the accuracy of exchange rate forecasts. In Zhang, et al. [10]a hybrid model was proposed to improve the accuracy of short-term electrical load forecasts. This model combines improved empirical mode decomposition, autoregressive integrated moving average, and wavelet neural networks, which are subsequently optimized by the fruit fly optimization algorithm (FOA). The results show that this hybrid model outperforms other models in terms of forecast accuracy.

Although there are several studies that address the forecasting of residential electric energy, there are no concrete studies of this nature applied to West African countries and more specifically to the city of Lomé in Togo. Additionally, hybrid methods developed to categorize types of residences in a given region have not yet been addressed. In the field of residential prediction, it is of utmost importance for producers or prosumers (simultaneous producers and consumers) of electricity to know the consumption of each residence at all times to better manage the recurring problem of the amount of energy produced, distribution, and consumption. Kalman filters and wavelet transform have the ability to process non-stationary signals and extract relevant information from noisy residential electrical energy data. This work addressed the forecasting of electrical energy for 28 residences collected in the city of Lomé using three hybrid models: WNNA, WNNK, and WKA, followed by the categorization of residences based on a compromise between the performance of each hybrid model and the characteristics of the residence data.

3. Materials and Methods

For this study, residential electrical energy data was collected every minute for 28 houses from a dataset gathered in Lomé, southern Togo, West Africa, covering the period from 15/07/2021 10:15 to 28/12/2023 23:59 for all 28 residences considered. The data collected from various locations in Lomé includes parameters such as active power, reactive power, apparent power, voltage, current, harmonics, and anomalies such as overvoltage, voltage imbalance, voltage dips, and load shedding. All these parameters and anomalies in the residential electrical network contribute to providing a clear picture of the electricity consumption profile of each residence. The collected data is decomposed into one-minute intervals selected using the PEL 106 data collection module during the collection process. This collected data provides information on load shedding frequencies, the nature of the energy source used in the residence, the quality of the electrical installation, and the amount of electrical energy consumed for each residence. Data collection is the only means to obtain the necessary data for studying and understanding the electricity consumption profile of each residence. This work can make electrical networks more dynamic and flexible for the development of a smarter system for integrating decentralized production sources by developing robust methods for forecasting residential electrical energy in Lomé.

3.1. Wavelet Method

Wavelets can adapt to the signal being processed with the ability to stretch or compress. These properties of compact function and null support function give them the capability to analyze and process unstable and noisy signals. The wavelet transforms functions as a signal adapter, acting on the signal to be analyzed by varying the scale of signal analysis to extract details present at different resolutions [6]. For a signal to be considered a wavelet, it must satisfy certain specific mathematical conditions: The first condition is that of admissibility, defined by equations 1 and 2.

$$C_\psi = \int_0^{+\infty} \frac{\left|\psi^{'}(\omega)\right|^2}{|\omega|} d\omega < +\infty \quad (1)$$

$$\psi^{'}(\omega) = \int_{-\infty}^{+\infty} \psi(t) e^{-i(2\pi f)t} dt \quad (2)$$

This means that the total energy of the wavelet must be finite. This condition ensures that the wavelet transform is invertible. Here the Fourier transform of $\psi(t)$ gives $\psi'(\omega)$and$c_\psi$is the

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

admissibility constant. The second condition is the zero-moment condition, which is defined by relation 3.

$$\int_{-\infty}^{+\infty} \psi(t)dt = \psi(f=0) = 0 \quad (3)$$

This means that the wavelet must have a zero mean. In other words, the continuous component of the wavelet must be removed, allowing the wavelet to capture local variations in the signal. The third condition is Localization, which states that the wavelet must be well localized in both the time domain and the frequency domain as explained in relation 4.

$$\psi_{x,y}(t) = \frac{1}{\sqrt{|x|}}\psi\left(\frac{t-y}{x}\right) \quad (4)$$

With x representing the scale parameter, which is used to compress or stretch the signal, and y representing the parameter that translates the temporal position of the wavelet along the signal. For a discrete wavelet decomposition of time series, relations 5 and 6 are respectively the scaling function and the mother wavelet function [20].

$$\psi_{j,k}(a) = \sqrt{2^j}\psi(2^j - a) \quad (5)$$

$$\varphi_{j,k}(a) = \sqrt{2^j}\varphi(2^j - a) \quad (6)$$

$AC_{J,k}$and $DC_{j,k}$are respectively the approximation and detail coefficients obtained after convolution of the mother wavelet function $\varphi_{J,k}$and scaling function $\psi_{j,k}$with the original signal f(t) as described in expressions 7 and 8.

$$AC_{J,k} = \int_{-\infty}^{+\infty} f(t)\varphi_{J,k}dt \quad (7)$$

$$DC_{j,k} = \int_{-\infty}^{+\infty} f(t)\psi_{j,k}dt \quad (8)$$

The residential electrical energy data considered is discretized and controlled over time. If we consider$y(t)$ as the time-discretized series whose length is given by the relation $k = 2^j$, equation 9 can be written.

$$y(t) = \sum_{k=-\infty}^{2^{J-k}-1} AC_{J,k}\varphi_{j,k}(t) + \sum_{j=1}^{J}\sum_{k=-\infty}^{2^{J-k}-1} DC_{j,k}\psi_{j,k}(t) \quad (9)$$

3.2. ARIMA Model

The ARIMA model is a statistical forecasting model that predicts future values of a time series by making a linear combination of past values. The ARIMA model is distinguished by three parts: the autoregressive (AR) part characterized by the parameter p, the differencing (I) part characterized by the parameter d, and the moving average (MA) part represented by the parameter q. p, d, and q are the hyperparameters that need to be found and used to obtain an optimal forecast. These three parameters are related by relation 10.

$$\left(1 - \sum_{k=1}^{p} \varphi_k L^k\right)(1-L)^d = \left(1 + \sum_{n=1}^{q} \theta_n L^n\right)\varepsilon_n \quad (10)$$

In this relationship,$\varepsilon_n$expresses the random error; $\varphi_k$and $\theta_n$are coefficients used in the ARIMA model. The Akaike Information Criterion (AIC) is used to select the optimal p, d, and q parameters for modeling the ARIMA model. This criterion is defined by relation 11.

$$AIC = 2\big(Nt - ln(Lv)\big) \quad (11)$$

With Nt and Lv representing respectively the total number of parameters to be used and the maximum value of the ARIMA model parameter function determined from the observed data. The p, d, q, parameters that correspond to the optimal model are those with the lowest Akaike Information Criterion (AIC) value. The AIC criterion is based on a logic that balances the complexity of the model used for forecasting and the efficiency of the fit.

3.3. Neural Networks Model

For neural network modeling, a given number of layers is defined, and each of these layers comprises a group of interconnected neurons. Thus, an input layer, an output layer, and an intermediate layer, which is a hidden layer, are associated. Neural networks use synaptic weights, which are stochastically initialized and adjusted to obtain an optimal prediction. The weighted inputs are added at each neuron, thus linking the activation function and the summation process [21]. This overall activation function clearly shows the non-linearity between the input system and the output system, defined as follows:

$$Y_m(t) = \sum_{n=1}^{N} \omega_{nm} x_n(t) + \beta_m \qquad (12)$$

This relationship ensures that activation functions, including logistic, ReLU, or hyperbolic functions, are all included to describe the nonlinear relationship between input and output. With Y_m ,ω_{mn},$\beta_m$and x_n representing the weighted sum, weights, bias, and input, respectively. Each neuron sums the weighted inputs and maps the summation to the output. Neural networks are widely used in the field of time series prediction using an iterative function that adjusts and corrects parameters at each iteration until optimal predicted values are obtained. Figure 3 shows the circuit from inputs to outputs.

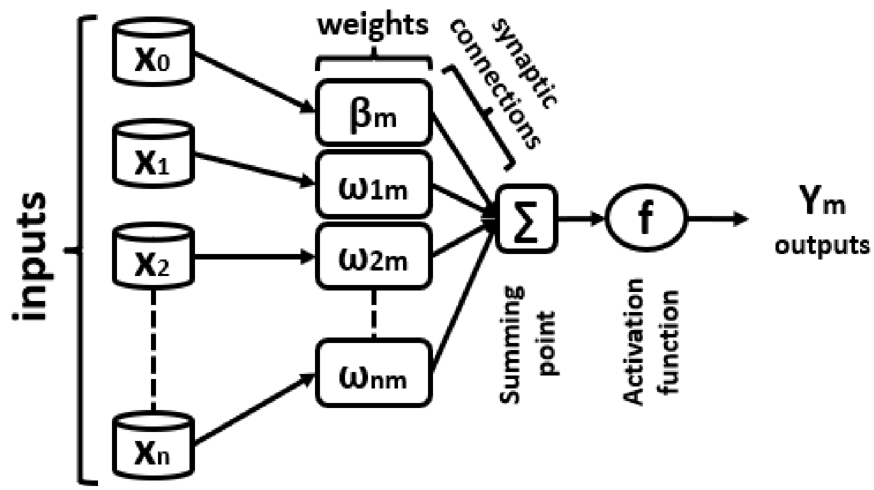

Figure1.
Architecture of Artificial Neurons.

3.4. The Kalman Filter

Relations 13 and 14 define the observed state equation of the algorithm used to model the processing by the Kalman filter [22].

$$\boldsymbol{x_t = \alpha x_{t-1} + \beta_t} \qquad (13)$$

$$y_t = \lambda x_t + \delta_t \qquad (14)$$

Both $x_t$and$y_t$are target state vectors. $x_t$is expressed as a function of the state transition matrix $\alpha$and the known state vector at time *t-1*, while $y_t$is estimated by the known state vector at time t with the state observation matrix λ. The matrices $\beta_t$and$\delta_t$represent the covariance of the system excitation, sometimes perceived as noise, and the covariance of the observation noise, respectively. The Kalman filter, based on mathematical and applied physics principles, allows the assimilation of the covariance of excitation noises of the entire system to be modeled. Relations 15 and 16, expressed by $\hat{x}_{t|t}$ and $\hat{x}_{t+1|t}$ show the evident recursion of the signal using the Kalman filter. These relations explicitly take into account past data to estimate the current state. The update of this estimation is done by considering the current state and past states to estimate a new state. $x_t$and Y_t being Gaussian functions, it is possible to use the following standard formulas to predict $\hat{x}_{t+1|t}$

$$\hat{x}_{t|t} = \bar{x}_t + \sum_{x_t y_t} \sum_{y_t}^{-1} (y_t - \overline{y}_t) \qquad (15)$$

$$\hat{x}_{t+1|t} = \alpha \hat{x}_{t|t} \qquad (16)$$

The Kalman filter will be used for preprocessing, which involves reducing noise in a signal, making real-time estimates, or being hybridized with another model to achieve better results. The algorithm used to model the Kalman filter follows the process represented in Figure 2.

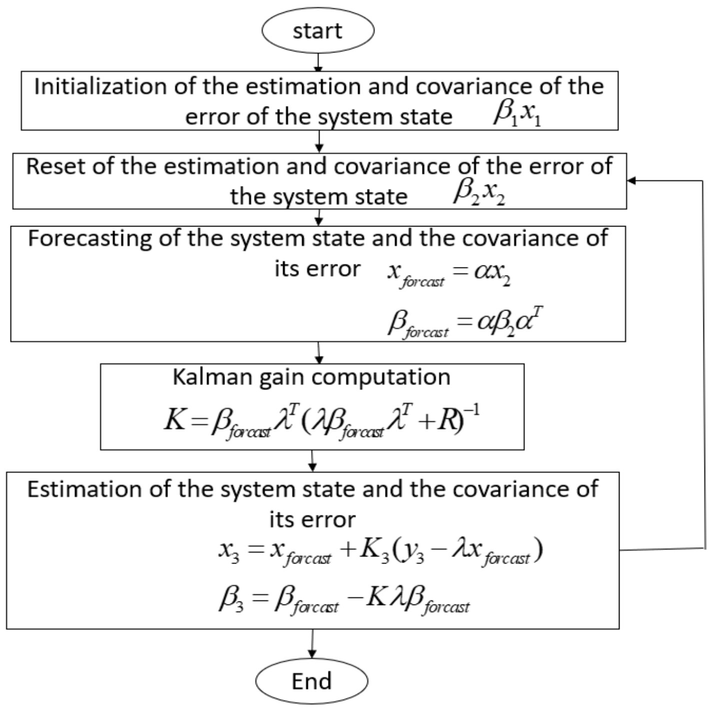

Figure 2.
The Kalman Filter Modeling Algorithm.

T denotes the timestamp for the estimation of the state matrix, R is provided by the radar with the timestamp indicating when the measurement took place, and t is the measurement covariance matrix.

3.5. Hybrid Model

The first step of this study is wavelet decomposition. This step involves decomposing into its various approximation and detail components and selecting those that will be injected into other forecasting models. The optimal decomposition level to obtain the best results corresponding to each of the 28 residences is determined by the following formula [23],

$$L_h = int[log(N_h)] \qquad (17)$$

Where L_h represents the maximum allowable decomposition level by the signal and $N_h$the signal length of the data collected per residence. Table 1 summarizes the 28 houses and their optimal allowable decomposition level.

ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

Table 1.
The level of decomposition achieved for each residence.

Ri	*R1*	*R2*	*R3*	*R4*	*R5*	*R6*	*R7*	*R8*	*R9*	*R10*	*R11*	*R12*	*R13*	*R14*	*R15*	*R16*	*R17*	*R18*	*R19*	*R20*	*R21*	*R22*	*R23*	*R24*	*R25*	*R26*	*R27*	*R28*
L_h	3	3	2	1	4	3	1	1	4	4	2	4	3	3	3	2	1	1	2	2	3	3	3	3	4	1	3	3

Figure 3 shows the order in which the different detail and approximation coefficients are obtained based on the decomposition level chosen for each residence and the correspondence between the decomposition level and the residences whose signal length verified by relation 17 gives the concerned decomposition level.

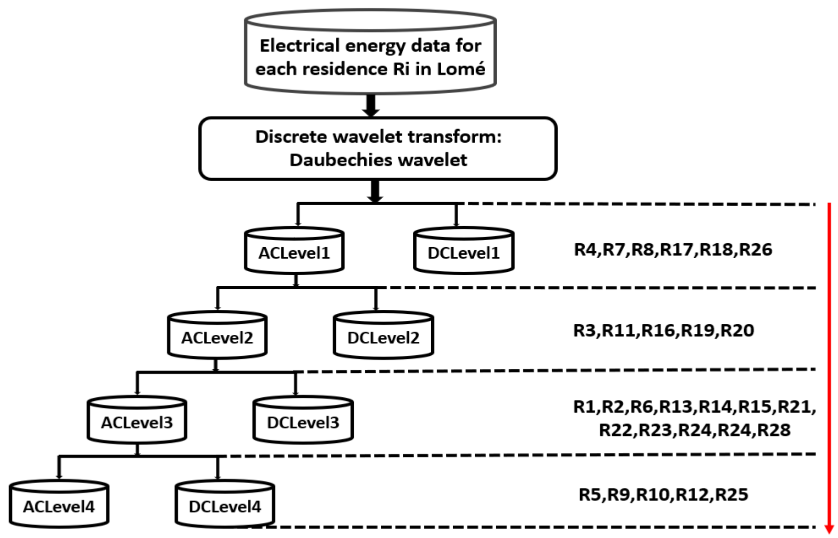

Figure 3.
Wavelet Decomposition of Different Residences.

Ri is the residence at rank i, with i ranging from 1 to 28. The next step is to use, according to the appropriate decomposition level, all the detail coefficients and the approximation coefficient of the last decomposition level. Table 2 clearly shows each level of decomposition and the different coefficients obtained.

Table 2.
Decomposition Levels and Obtained Coefficients.

Decomposition level	**Coefficients obtained**
Level 1	ACLevel1+DCLevel1
Level 2	ACLevel2+DCLevel2+DCLevel1
Level 3	ACLevel3+DCLevel3+DCLevel2+DCLevel1
Level 4	ACLevel4+DCLevel4+DCLevel3+DCLevel2+DCLevel1

Daubechies wavelets of order 3 are chosen as the mother wavelet due to their optimal adaptation and their ability to more easily identify the information carried by residential electrical load signals [24].

ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

3.5.1. Wavelet Coupled to Neural Network and ARIMA (WNNA)

The hybrid wavelet-neural network-ARIMA model consists of several steps. The first step involves performing discrete wavelet decomposition. After obtaining the various coefficients from the decomposition, the approximation coefficient from the last level of decomposition and all the detail coefficients obtained from the first level to the last level of decomposition are retained for further modeling. The second step involves using the approximation coefficient (ACLevel n) to inject it into the ARIMA model and the detail coefficients (DCLevel n, DCLevel n-1, ..., DCLevel 1) to inject them separately into the neural network model for prediction purposes (n represents the maximum decomposition level of the electrical energy data of residence Ri). The final step involves predicting the different coefficients and summing these predicted coefficients to obtain the final prediction. Figure 4 represents the flowchart followed to model the hybrid WNNA model.

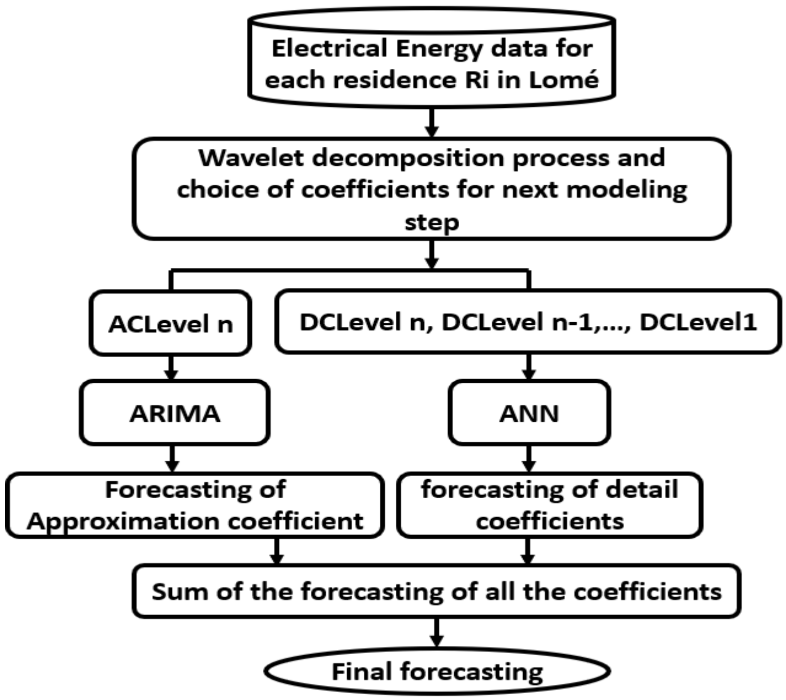

Figure 4.
Diagram characterizing the WNNA model.

Due to the fact that the approximation parts are low-frequency components showing the signal trend, ACLevel n will be modeled with the ARIMA model. The detail parts, on the other hand, represent high-frequency components that contain a lot of undesirable noise. This noise is preprocessed using denoising techniques. After this preprocessing, the detail coefficients are less noisy but exhibit complexity and non-linearity. Therefore, the neural network model is used to model the coefficients DCLevel n, DCLevel n-1, ..., DCLevel 1.

3.5.2. Wavelet Coupled to Neural Network and Kalman (WNNK)

The hybrid model proposed in this section is based on wavelet decomposition, the Kalman filter, and neural networks. The approximation and detail coefficients are separately introduced into the Kalman filter algorithm, which eliminates interferences and reduces noise on each coefficient. After processing by the Kalman filter, we obtain the estimates of each coefficient. At the input, the coefficients ACLevel n(i-1), DCLevel n(i-1), ACLevel n-1(i-1), ..., DCLevel 1(i-1) become, respectively, at the output of the Kalman filter, ACLevel n(i), DCLevel n(i), ACLevel n-1(i), ..., DCLevel 1(i). ei represents the set of these

ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

coefficients obtained at the output of the Kalman filter, so e1 would be equal to the coefficient ACLevel n(1), DCLevel n(1), ACLevel n-1(1), ..., DCLevel 1(1), considered separately when injected into the neural network model. e2, in turn, will be equal to ACLevel n(2), DCLevel n(2), ACLevel n-1(2), ..., DCLevel 1(2), and so on up to ei. Yi represents the output of the neural networks where the sum of the different coefficient estimates is already applied.

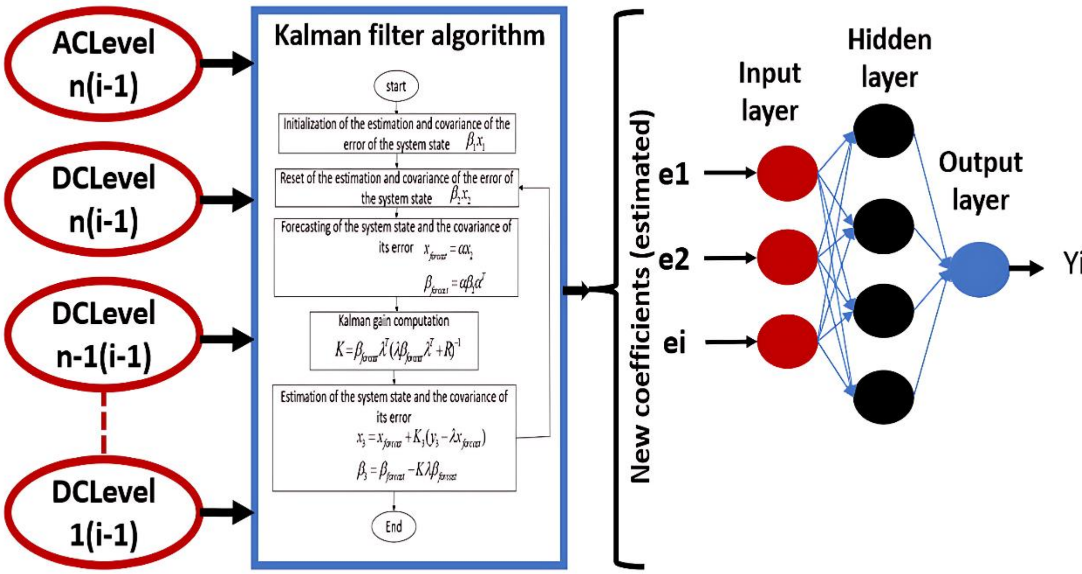

Figure 5.
Outlines the structure of the hybrid WNNK model.

This method, identified as a robust method due to the non-linearity of electrical energy data, is applied to the entire dataset of each residence for forecasting.

3.5.3. Wavelet Coupled to ARIMA and Kalman (WKA)

This method, based on discrete wavelet decomposition, the Kalman filter, and the ARIMA model, combines the advantages of wavelets for signal decomposition, the Kalman filter for state estimation of each coefficient obtained after decomposition, and ARIMA for final modeling. After the final prediction of the different coefficients, we obtain Yi, which represents the sum of the predictions of the different coefficients at the output of the ARIMA model. Figure 6 shows the structural diagram of the hybrid WKA method.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

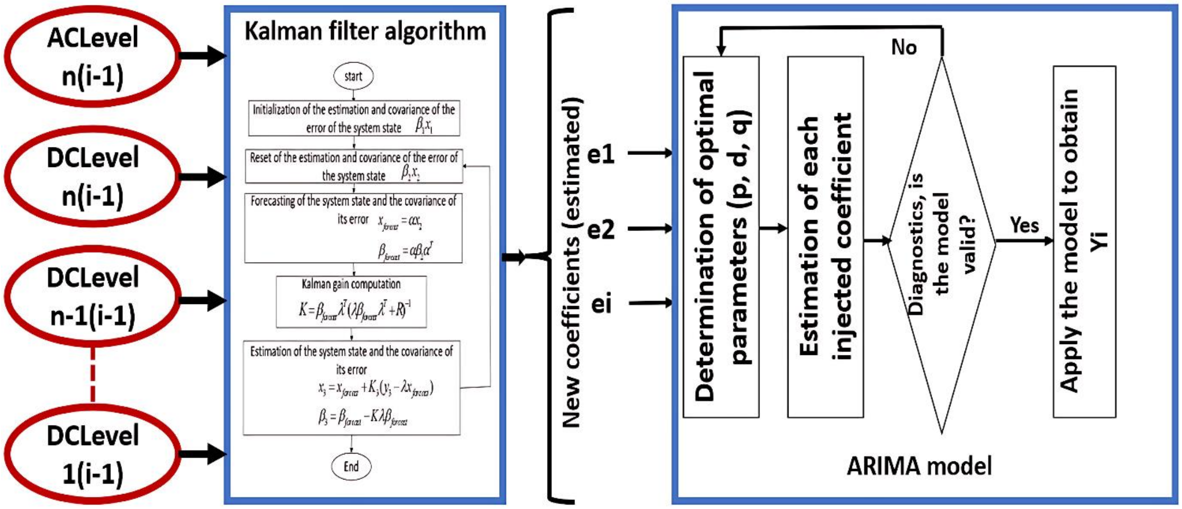

Figure 6.
Structural Diagram of the Hybrid WKA Method.

4. Evaluation Metrics

The Root Mean Square Error (RMSE) gives an idea of the dispersion of residuals. Lower RMSE values indicate a better fit of the model to the data. The formula that determines the RMSE is defined by relation 18 [25],

$$RMSE = \sqrt{\frac{1}{n}\sum_{t-1}^{n}[x(t) - x'(t)]^2} \tag{18}$$

With n representing the total number of observations, $x(t)$ the observed values, and $x'(t)$ the values predicted by the model.

The Mean Absolute Percentage Error (MAPE) provides a clear interpretation of the magnitude of the error in percentage terms, with lower values indicating better model performance. Relation 19 allows for the calculation of the MAPE [27].

$$MAPE = \frac{1}{n}\sum_{t-1}^{n}\left|\frac{x(t)-x'(t)}{x(t)}\right| \times 100 \tag{19}$$

With n representing the total number of observations, $x(t)$ the observed values, and $x'(t)$ the values predicted by the model.

The correlation coefficient (R^2) ranges between 0 and 1, where 0 indicates no correlation, and 1 indicates a strong correlation. Relation 20 allows for the calculation of the R^2 [26],

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x(t)-x'(t))^2}{\sum_{i=1}^{n}((x(t)-\overline{x})^2} \tag{20}$$

With n representing the total number of observations, $x(t)$ the observed values, $x'(t)$ the values predicted by the model and $\overline{x}$ is the mean of the observed values. These criteria allow the evaluation of the accuracy and robustness of forecasting models.

5. Discussion of the Results

The different models were developed and applied to the electrical energy datasets of each of the 28 residences for prediction purposes. The performances provided by the different hybrid models are compared to each other in order to find accurate and relevant prediction results with the least possible error for each residence. Table 3 shows the predictive performance of the 3 hybrid models on the 28 residences.

Table 3.
Predictive Performance of Hybrid Models on the 28 Residences.

Model	**WNNA**			**WNNK**			**WKA**		
Error / Ri	**RMSE**	**MAPE**	**R^2**	**RMSE**	**MAPE**	**R^2**	**RMSE**	**MAPE**	**R^2**
R1	0.3085	0.1089	0.8072	0.0641	0.0245	0.9701	0.3482	0.1826	0.7110
R2	0.1925	0.0753	0.9041	0.2443	0.1901	0.7803	0.4754	0.2230	0.6891
R3	0.1791	0.0739	0.9109	0.2812	0.2011	0.7761	0.4492	0.2152	0.6594
R4	0.5201	0.3011	0.7317	0.4862	0.3071	0.7372	0.1896	0.0856	0.9381
R5	0.4758	0.1820	0.7901	0.0547	0.0482	0.9785	0.4207	0.1918	0.7978
R6	0.2982	0.1584	0.8042	0.0426	0.0251	0.9683	0.3821	0.2205	0.7273
R7	0.3841	0.1825	0.8145	0.0487	0.0348	0.9706	0.4895	0.3308	0.6842
R8	0.4824	0.2847	0.7457	0.4251	0.2958	0.7390	0.1564	0.0767	0.9521
R9	0.0987	0.0225	0.9582	0.1897	0.1457	0.8285	0.3933	0.2014	0.7099
R10	0.4257	0.3044	0.8294	0.4335	0.2484	0.7580	0.1875	0.0905	0.9596
R11	0.4982	0.2034	0.7524	0.0687	0.0342	0.9591	0.4987	0.2624	0.7604
R12	0.5081	0.2351	0.7643	0.3987	0.2384	0.7492	0.1753	0.0568	0.9621
R13	0.5279	0.3728	0.8631	0.5039	0.2874	0.7561	0.0982	0.0428	0.9729
R14	0.1074	0.0532	0.9281	0.2089	0.2008	0.8021	0.4268	0.2601	0.6986
R15	0.5012	0.2881	0.7208	0.0385	0.0215	0.9705	0.4821	0.2035	0.7058
R16	0.5174	0.2396	0.8025	0.4832	0.2147	0.7421	0.0849	0.0628	0.9782
R17	0.4842	0.2451	0.6924	0.0814	0.0254	0.9681	0.4521	0.2564	0.7105
R18	0.1904	0.0820	0.9082	0.2552	0.1992	0.7928	0.4356	0.2504	0.7109
R19	0.1891	0.0864	0.9318	0.2406	0.1877	0.8083	0.4805	0.2620	0.8012
R20	0.4369	0.3046	0.7613	0.4157	0.2541	0.7857	0.0793	0.0845	0.9765
R21	0.4631	0.3007	0.7922	0.4554	0.2754	0.7658	0.1092	0.0958	0.9699
R22	0.5170	0.2901	0.6881	0.0753	0.0267	0.9758	0.4952	0.2387	0.6899
R23	0.4129	0.2642	0.7733	0.4895	0.2814	0.8089	0.0888	0.0927	0.9795
R24	0.0991	0.0427	0.9428	0.1908	0.1562	0.8108	0.4042	0.2115	0.7509
R25	0.4935	0.2140	0.8113	0.4458	0.2357	0.7892	0.1982	0.0687	0.9808
R26	0.1729	0.0842	0.9349	0.2256	0.2101	0.8176	0.4385	0.2780	0.6991
R27	0.4856	0.1992	0.7345	0.0895	0.0354	0.9752	0.4951	0.1572	0.8056
R28	0.4231	0.2123	0.8325	0.0971	0.0481	0.9810	0.4824	0.1842	0.8149

This table shows that it is possible to classify the residences into 3 groups based on the prediction results. Group 1 consists of 10 residences (R1, R5, R6, R7, R11, R15, R17, R22, R27, R28) which give the best results with the hybrid WNNK model. Group 2 includes 8 residences (R2, R3, R9, R14, R18, R19, R24, R26) and gives the best results with the hybrid WNNA model. Group 3 consists of 10 residences (R4, R8, R10, R12, R13, R16, R20, R21, R23, R25) which give the best results with the hybrid WKA model. The error evaluation criteria applied to Group 1 show R2 values ranging from 0.9591 to 0.9810 for the hybrid WNNK model, while the other hybrid models for the same group give R2 values ranging from 0.6881 to 0.8325 for the WNNA model and from 0.6842 to 0.8149 for the WKA model. For Group 2, the R2 values range from 0.9041 to 0.9582 for the hybrid WNNA model, while the other hybrid models for the same group give R2 values ranging from 0.7761 to 0.8285 for the WNNK model and from 0.6594 to 0.8012 for the WKA model. Finally, for Group 3, the R2 values range from 0.9381 to 0.9808 for the hybrid WKA model, while the other hybrid models for the same group give R2 values ranging from 0.7317 to 0.8631 for the WNNA model and from 0.7372 to 0.8089 for the WNNK model. Three residences, each belonging to one of the three groups, are visualized in the following figures, highlighting the predictions given by the three hybrid models each time. Based on the residences that give the highest correlation coefficients, the choice for visualization was made for residence R28 for Group 1, R9 for Group 2, and R25 for Group 3.

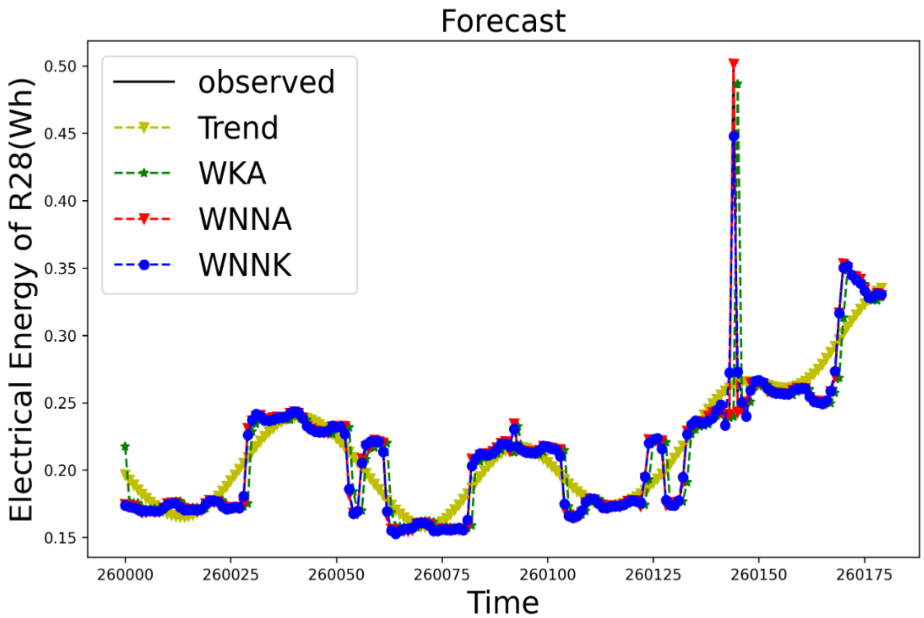

Figure 7.
Comparison of Different Results of Hybrid Models for Residence R28.

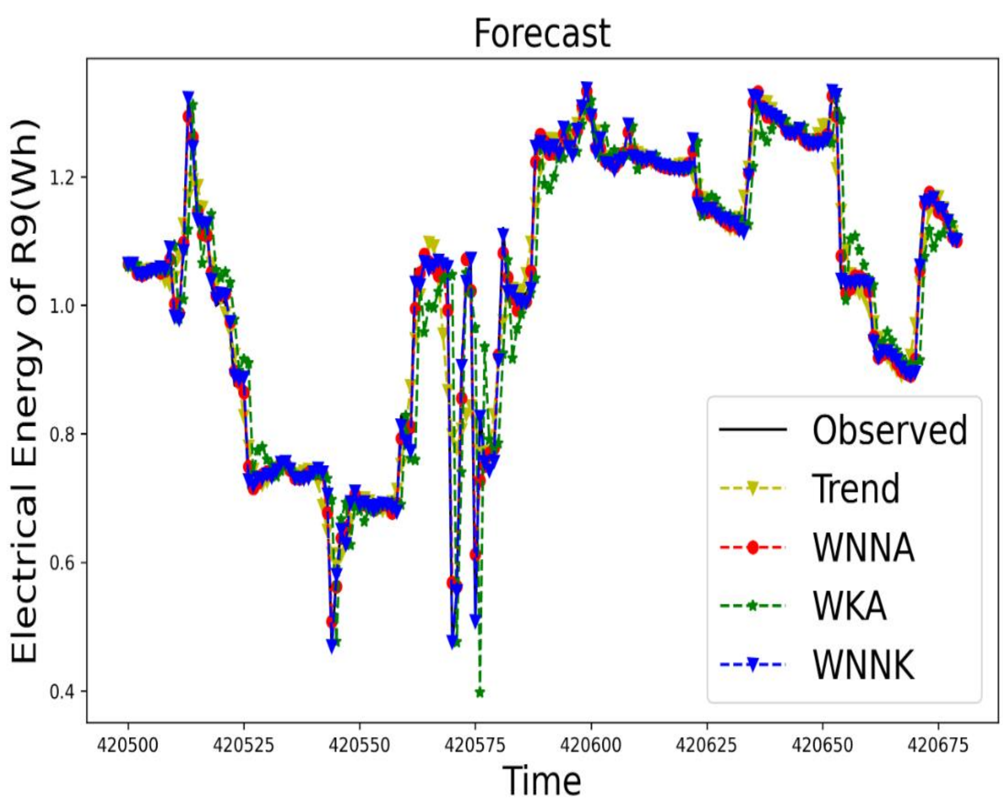

Figure 8.
Comparison of Different Results of Hybrid Models for Residence R9.

ISSN: 2576-8484
Vol. 9, No. 3: 1892-1907, 2024
DOI: 10.55214/25768484.v9i3.5716

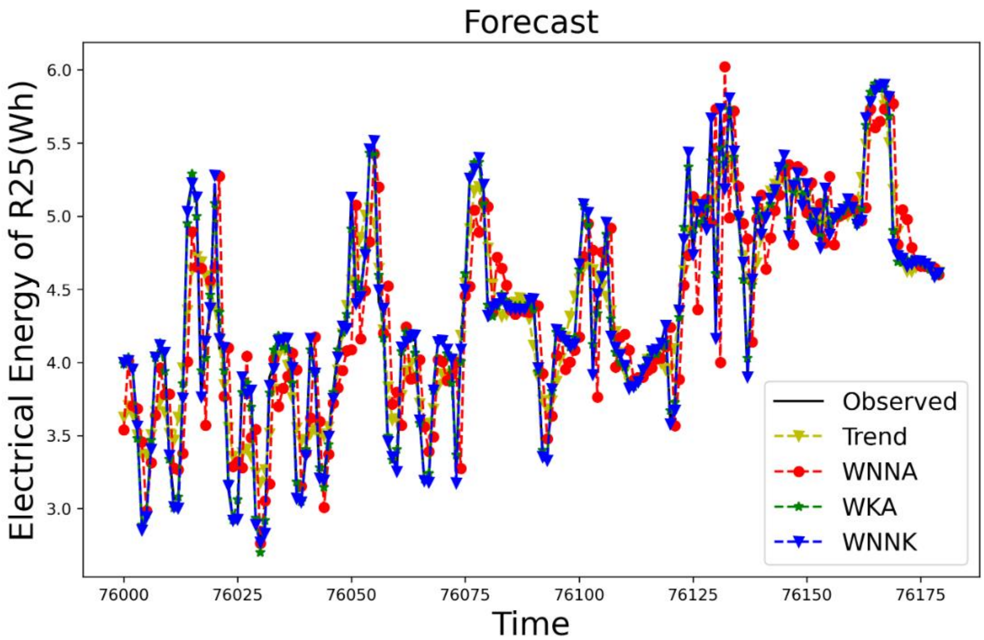

Figure 9.
Comparison of Different Results of Hybrid Models for Residence R25.

These graphs in Figures 7, 8, and 9 show the predicted values by the hybrid models WNNA, WNNK, WKA and the actual observed values used to test the models for the considered residences. As shown in Figure 7, the prediction given by residence 28 of Group 1 with the hybrid WNNA model is more correlated with the observed signal than with the WNNK and WKA models. Similarly, in Figure 8, the prediction of residence 9 of Group 2 given by the hybrid WNNK model shows a better correlation with the actual observed signal than that given by the WNNA and WKA models. The prediction of residence 25 of Group 3 given by the hybrid WKA model has a stronger correlation with the actual observed signal than that given by the WNNA and WNNK models. According to the obtained results, the signal trend provides information about the nature of the data, showing how unstable the signal is. The smoother the signal and the lower the amplitude (between 0 and 1watt-hour (Wh)), the better the results with the hybrid WNNK model. For trends where the signal is very unstable and noisy with many fluctuations, there are two possibilities: one with very low amplitudes (between 0 and 2 Wh) which offers the best results with the hybrid WNNA model, and the other with relatively high amplitudes (between 0 and 7 Wh) which gives better results with the hybrid WKA model. The data of the residences belonging to a given group present several similarities such as the level of electrical voltage, the intensity of the electrical current, the network frequency, the consumed and produced power (active, reactive, and apparent), the quality of harmonics, voltage dips and peaks, the quality of the power factor, event capture (load shedding), the nature of the energy source used in the residence, the quality of the electrical installation performed, and the average consumed power.

6. Conclusions

Mastering the consumption habits of residences requires a thorough analysis of the collected data followed by accurate predictions for each residence. This is the foundation of the dynamism of the residential energy sector, enabling optimal decision-making by the network manager and the potential integration of decentralized production sources when necessary. In this work, three hybrid models were developed based on wavelet decomposition, which is then combined with neural networks and the Kalman filter, or neural networks and ARIMA, or the Kalman filter and ARIMA, to find a suitable prediction model for each category of residence. The analysis of the collected electrical energy data allows classifying a given residence into its respective group. As each group is affiliated with a hybrid model, error verification tests performed with RMSE, MAPE, and R2 have confirmed that residences belonging to Group 1, Group 2, and Group 3 respectively provide optimal prediction results with the hybrid models WNNK, WNNA, and WKA. Mastering electrical energy at the residential level contributes to accurately understanding the consumption pattern of each residence. By categorizing each residence into one of these three groups, it is possible to determine in advance the appropriate hybrid prediction model for each residence. This mastery of energy at the residential level will be beneficial not only to consumers but also to the electrical network manager by offering the possibility to anticipate future electrical energy consumption and take necessary measures for the potential integration of decentralized production sources.

Funding:
This work was financially supported by the World Bank through CERME.

Institutional Review Board Statement:
Not applicable.

Transparency:
The authors confirm that the manuscript is an honest, accurate and transparent account of the study that no vital features of the study have been omitted and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Competing Interests:
The authors declare that they have no competing interests.

Authors' Contributions:
Conceptualization, Adekunlé Akim Salami and Yéloiti N'poh Babikou; Methodology, Adekunlé Akim Salami; Software, Yéloiti N'poh Babikou; Validation, Adekunlé Akim Salami and Koffi Mawugno Kodjo; Formal analysis, Agbassou Guenoukpati; Investigation, Yéloiti N'poh Babikou; Resources, Adekunlé Akim Salami; Data conservation, Yéloiti N'poh Babikou; Writing - Original Draft preparation, Yéloiti N'poh Babikou; Writing - Reviewing and editing, Adekunlé Akim Salami and Agbassou Guenoukpati; Visualization, Yéloiti N'poh Babikou; Supervision, Adekunlé Akim Salami and Koffi Mawugno Kodjo.

Acknowledgements:
The authors of this paper would like to express their sincere gratitude to the World Bank for funding this research through CERME (Centre d'excellence Régional pour la Maîtrise de l'Electricité) at the University of Lomé.

References
[1] S. Islam and N. K. Roy, "Renewables integration into power systems through intelligent techniques: Implementation procedures, key features, and performance evaluation," *Energy Reports*, vol. 9, pp. 6063-6087, 2023. https://doi.org/10.1016/j.egyr.2023.05.063

[2] A. R. Singh, R. S. Kumar, M. Bajaj, C. B. Khadse, and I. Zaitsev, "Machine learning-based energy management and power forecasting in grid-connected microgrids with multiple distributed energy sources," *Scientific Reports*, vol. 14, no. 1, p. 19207, 2024. https://doi.org/10.1038/s41598-024-70336-3

[3] M. Pullinger *et al.*, "The IDEAL household energy dataset, electricity, gas, contextual sensor data and survey data for 255 UK homes," *Scientific Data*, vol. 8, no. 1, p. 146, 2021. https://doi.org/10.1038/s41597-021-00921-y

[4] M. Schlemminger, T. Ohrdes, E. Schneider, and M. Knoop, "Dataset on electrical single-family house and heat pump load profiles in Germany," *Scientific Data*, vol. 9, no. 1, p. 56, 2022. https://doi.org/10.1038/s41597-022-01156-1

[5] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72-81, 2019. https://doi.org/10.1016/j.energy.2019.05.230

[6] P. K. Dharmoju, K. Yeluripati, J. Guduri, and K. Palle, "Forecasting electrical demand for the residential sector at the national level using deep learning," in *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, 2021: IEEE, pp. 1-6.

[7] Z. A. Khan, A. Ullah, W. Ullah, S. Rho, M. Lee, and S. W. Baik, "Electrical energy prediction in residential buildings for short-term horizons using hybrid deep learning strategy," *Applied Sciences*, vol. 10, no. 23, p. 8634, 2020. https://doi.org/10.3390/app10238634

[8] X. Lin, R. Zamora, C. A. Baguley, and A. K. Srivastava, "A hybrid short-term load forecasting approach for individual residential customer," *IEEE Transactions on Power Delivery*, vol. 38, no. 1, pp. 26-37, 2022. https://doi.org/10.1109/TPWRD.2022.3178822

[9] S. Sulaiman, P. A. Jeyanthy, D. Devaraj, and K. Shihabudheen, "A novel hybrid short-term electricity forecasting technique for residential loads using Empirical Mode Decomposition and Extreme Learning Machines," *Computers & Electrical Engineering*, vol. 98, p. 107663, 2022. https://doi.org/10.1016/j.compeleceng.2021.107663

[10] J. Zhang, Y.-M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, pp. 774-781, 2018. https://doi.org/10.1016/j.energy.2018.06.012

[11] H. H. Aly, "A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid," *Electric Power Systems Research*, vol. 182, p. 106191, 2020. https://doi.org/10.1016/j.epsr.2019.106191

[12] Z. Zheng, H. Chen, and X. Luo, "A Kalman filter-based bottom-up approach for household short-term load forecast," *Applied Energy*, vol. 250, pp. 882-894, 2019. https://doi.org/10.1016/j.apenergy.2019.05.102

[13] C. Guan, P. B. Luh, L. D. Michel, and Z. Chi, "Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3806-3817, 2013. https://doi.org/10.1109/TPWRS.2013.2264488

[14] P. V. B. Ramos, S. M. Villela, W. N. Silva, and B. H. Dias, "Residential energy consumption forecasting using deep learning models," *Applied Energy*, vol. 350, p. 121705, 2023. https://doi.org/10.1016/j.apenergy.2023.121705

[15] G. La Tona, M. Luna, and M. Di Piazza, "Day-ahead forecasting of residential electric power consumption for energy management using Long Short-Term Memory encoder–decoder model," *Mathematics and Computers in Simulation*, vol. 224, pp. 63-75, 2024. https://doi.org/10.1016/j.matcom.2023.06.017

[16] G.-F. Fan, Y. Zheng, W.-J. Gao, L.-L. Peng, Y.-H. Yeh, and W.-C. Hong, "Forecasting residential electricity consumption using the novel hybrid model," *Energy and Buildings*, vol. 290, p. 113085, 2023. https://doi.org/10.1016/j.enbuild.2023.113085

[17] F. Gao, H. Chi, and X. Shao, "Forecasting residential electricity consumption using a hybrid machine learning model with online search data," *Applied Energy*, vol. 300, p. 117393, 2021. https://doi.org/10.1016/j.apenergy.2021.117393

[18] Y. Patel and D. Deb, "Machine intelligent hybrid methods based on Kalman filter and wavelet transform for short-term wind speed prediction," *Wind*, vol. 2, no. 1, pp. 37-50, 2022. https://doi.org/10.3390/wind2010003

[19] M. Khashei and B. Mahdavi Sharif, "A Kalman filter-based hybridization model of statistical and intelligent approaches for exchange rate forecasting," *Journal of Modelling in Management*, vol. 16, no. 2, pp. 579-601, 2021. https://doi.org/10.1108/JM2-12-2019-0277

[20] S. Singh, K. S. Parmar, J. Kumar, and S. J. S. Makkhan, "Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19," *Chaos, Solitons & Fractals*, vol. 135, p. 109866, 2020. https://doi.org/10.1016/j.chaos.2020.109866

[21] A. Di Piazza, M. C. Di Piazza, G. La Tona, and M. Luna, "An artificial neural network-based forecasting model of energy-related time series for electrical grid management," *Mathematics and Computers in Simulation*, vol. 184, pp. 294-305, 2021. https://doi.org/10.1016/j.matcom.2020.05.010

[22] Y. Yang, T. Yu, W. Zhao, and X. Zhu, "Kalman filter photovoltaic power prediction model based on forecasting experience," *Frontiers in Energy Research*, vol. 9, p. 682852, 2021. https://doi.org/10.3389/fenrg.2021.682852

[23] Y. Wang, H. Cheng, H. Chen, M. Ye, Y. Ren, and C. Yang, "A hybrid model based on wavelet decomposition and LSTM for short-term energy consumption prediction of chillers," *Journal of Building Engineering*, vol. 99, p. 111539, 2025. https://doi.org/10.1016/j.jobe.2024.111539

[24] S. Shiralashetti, "An application of the daubechies orthogonal wavelets in power system engineering," *Int. J. Comput. Appl*, vol. 975, p. 8878, 2014.

[25] A. Guenoukpati, A. Pierre, K. Amussou, A. Akim, and S. Ouedraogo, "Optimal power density estimation in Sub-Saharan Africa regions based on Weibull distribution parameters," *Edelweiss Applied Science and Technology*, vol. 8, no. 4, pp. 1937-1955, 2024. https://doi.org/10.55214/25768484.v8i4.1568

[26] M. S. Salim, N. Sabri, and A. A. R. Dheyab, "Artificial neural network weights-based new formula for solar power plant energy prediction," *Edelweiss Applied Science and Technology*, vol. 8, no. 6, pp. 9345-9355, 2024. https://doi.org/10.55214/25768484.v8i6.3998