

Article

High-Accuracy Energy Forecasting for Sustainable Hospitality: A Hybrid Ensemble Machine Learning Approach to 50-Year Retrofit Analysis in Sub-Tropical Hotels

Milen Balbis-Morejón ¹, Oskar Cabello-Justafre ², Juan José Cabello-Eras ³, Javier M. Rey Hernández ^{4,5}, Francisco J. Rey-Martínez ^{5,6}, A. O. Elgharib ⁷ and Khaled M. Salem ^{5,7,*}

¹ Department of Energy, Universidad de la Costa, Calle 58 No. 55-66, Barranquilla 080002, Colombia; mbalbis1@cuc.edu.co

² Departament of Mechanical Engineering, Universidad de Cienfuegos, Carretera a Rodas, km 4, Cienfuegos 59430, CS, Cuba; oskarcabellojustafre@gmail.com

³ Departament of Mechanical Engineering, Universidad de Córdoba, Carrera 6 No. 77-305, Montería 230002, Colombia; juancabelloe@correo.unicordoba.edu.co

⁴ Department of Mechanical Engineering, Fluid Mechanics and Thermal Engines, Engineering School, University of Malaga (UMA), 29016 Malaga, Spain; jrey@uma.es

⁵ Department of Energy and Fluid Mechanics, Engineering School (EII), University of Valladolid, Paseo del Cauce 59, 47011 Valladolid, Spain; rey@uva.es

⁶ GIRTER Research Group, Consolidated Research Unit (UIC053) of Castile and Leon, 47011 Valladolid, Spain

⁷ Department of Basic and Applied Science Engineering, Arab Academy for Science, Technology and Maritime Transport (Smart Village Campus), Smart Village, Giza 12577, Egypt; ahmedomarhmed@aast.edu

* Correspondence: khaled.khaled24@estudiantes.uva.es

Abstract

Accurate energy forecasting is critical for the financial and environmental sustainability of the hospitality sector, particularly in energy-intensive subtropical climates. This research addresses a significant gap by conducting a comprehensive, comparative analysis of six machine learning algorithms—Artificial Neural Networks (ANN), Random Forest (RF), XGBoost, Radial Basis Function (RBF), Autoencoder, and Decision Trees—to predict the hourly energy consumption of a hotel in Cuba. We significantly enhance predictive performance through a novel hybrid ensemble scheme, integrating voting, stacking, and blending techniques. Furthermore, this study pioneers a long-term forecasting methodology by utilizing a Long Short-Term Memory (LSTM) model to project the hotel's energy demand over a 50-year horizon, providing the strategic insight necessary for evaluating major retrofits. Our results demonstrate that ensemble methods, particularly blending, achieve superior accuracy and stability, with correlation coefficients up to 0.975 and the lowest error metrics. The subsequent high-fidelity predictions, including an analysis revealing a minimal specific CO₂ emission of 0.025 kg from natural gas use, provide a quantitative foundation for formulating sustainable energy policies, incentivizing investment in efficient technologies, and ensuring that long-term emission reduction targets are both financially viable and technically robust.



Academic Editor: Seung-Hoon Yoo

Received: 21 January 2026

Revised: 16 February 2026

Accepted: 25 February 2026

Published: 27 February 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: climate change; AI; big data; energy forecasting; CO₂ emissions; nZEB; ensemble methods

1. Introduction

The global hospitality sector is facing increasing pressure to address its significant environmental footprint, driven primarily by higher operational energy demands. This problem

is particularly severe in subtropical and tropical climates, areas defined by persistently high temperatures and humidity. In such settings, the primary challenge for hotel operators is to keep guests comfortable and as a result, space cooling and dehumidification are the dominant contributors to energy use. Therefore heating, ventilation and air conditioning (HVAC) systems are typically the largest source of cost and environmental impact in hotel operations [1,2]. Retrofits are key to building modernization and competitiveness with a large share of the global hotel stock aging. Such renovations offer a significant, albeit all too often overlooked, opportunity to build in sustainability for the long term. Regrettably, plenty of retrofit work takes a piecemeal or shallow-working approach, concentrating on esthetics or small utility upgrades rather than whole-system energy solutions [3–5].

The development and brisk growth of artificial intelligence (AI), which is represented by intricate large language models (LLM) and generative AI, constitutes a basic and intensifying danger to energy sustainability on a global scale. The major player in this sector is the data center that is constructed in a way that consumes a colossal amount of power and is packed with rows of servers which are responsible for the training of models and delivering real-time responses to users (inference). The making of a single state-of-the-art model involves huge energy consumption, and the energy needs continue with the users' queries—from asking a chatbot for a summary to generating a unique image—that also require power, and the amount is often even more than that of traditional internet activities [6–8]. This appetite for computation is quickly becoming a primary source of growth in global electricity consumption, straining existing electricity grids and in some cases leading to an increase in fossil fuel-based generation to supply the immediate energy needs of these burgeoning technology hubs. But AI and energy have a nuanced relationship, and there is great potential for mitigation. Although it is the shortcomings of AI tech that are increasing the carbon footprint of this industry, they also have great potential as a tool for energy efficiency in different industries. AI tools could also be used to design “smarter, more energy-efficient electric grids” that better incorporate and anticipate intermittent renewable energy sources such as wind and solar [9,10].

1.1. Literature Review

This study by Yalcintas and Akkurt [11] successfully applies Artificial Neural Networks (ANN) to predict building energy use, focusing specifically on chiller plant consumption in tropical climates. They demonstrate that ANN models are advantageous due to their speed and adaptability in predicting energy savings from building retrofits. The developed model showed high accuracy, with a testing error of just 10.0%, confirming ANNs as a reliable tool for real-world energy analysis. S.L. Wong et al. [12] developed an accurate artificial neural network (ANN) model to predict daily electricity usage of daylighted office buildings in subtropical climates. The model used nine inputs, including weather, building design parameters, and type of day. Trained using EnergyPlus simulation data, the ANN demonstrated excellent prediction accuracy, with Nash-Sutcliffe efficiency coefficients close to 0.99 for cooling, lighting, and total energy use, confirming its reliability for energy analysis. Yuo-Hsien et al. [13] An artificial neural network (ANN) was used to successfully model the complex relationship of industrial energy demand, manufacturing output, and climate change in Taiwan. The improved ANN model revealed that industrial electricity demand is price inelastic and positively affected by climate change.

Muhammad et al. [14] utilized techniques of machine learning to make predictions on useful solar thermal energy (USTE) and amongst different models, it was revealed that Random Forest (RF) and Extra Trees (ET) offered the highest prediction accuracy. Although Decision Tree (DT) was the quickest to deliver results, the higher accuracy and stability

of RF and ET make them more suited for decentralized energy management and fault detection applications. Wang Zeyu et al. [15] conducted a study where they introduced Random Forest (RF) as an ensemble method for predicting hourly energy consumption in two educational buildings and the results were in favor of RF compared to Regression Trees (RT) and Support Vector Regression (SVR). The authors also suggested that the accuracy of RF models could be increased if the training data is monthly (or semester-specific), thus taking the changing operational behavior of the buildings into account. Suhua et al. [16] conducted a comprehensive global study using data from 237 Fluxnet towers to investigate and scale light-use efficiency (LUE), a key component in estimating global gross primary production (GPP). They employed a Random Forest Regression (RFR) approach to integrate tower-based LUE estimates with key environmental and biological variables at $0.5^\circ \times 0.5^\circ$ resolution, subsequently developing an RFR-LUE-GPP model.

In Spain, Salem et al. [17] tackled the issue of energy demand forecasting throughout three different building constructions by means of a comparison of six ML models, among them Random Forest, XGBoost, and Artificial Neural Networks (ANN). In the study performed by Salem et al. [18], various ML models were compared in terms of their effectiveness in predicting the hourly energy difference in consumption in different parts of Europe as a measure to facilitate the shift towards renewable energy sources. Salem et al. [1] studied the performance of six different AI models (ANN, RF), the main objective was to determine the best models for emissions forecasting in particular building situations, thereby filling the research void in comparative CO₂ forecasting across different buildings and climates. The analysis was further advanced using linear regression to forecast CO₂ emissions over the next three years, providing stakeholders with long-term insights to help improve energy efficiency and reduce the carbon footprint of the built environment [18–22]. Yanchao et al. [23] investigated the correlation between AI adoption and Carbon Emissions (CEs) in Chinese A-share listed companies (2009–2021). Their panel fixed-effects regression found that AI adoption has a significant negative impact on CEs, robust across various controls. The CE suppression effect is greater in non-polluting, non-high-tech, and capital-intensive industries, and regions with non-state-owned firms or strict environmental regulations. Bai et al. highlighted [24] that energy structure transition is a vital, long-term evolution toward low-carbon, clean, and reliable energy. This transition is essential for achieving national carbon reduction targets and ensuring the sustainable development of human societies. It necessitates developing a scientific approach to assess its multidimensional benefits (economic, social, and environmental). Jiahui et al. [25] found that by 2050, ecological overload will heavily impact lower-income nations, with the poorest 50% of countries bearing 70% of the burden.

Fathi et al. [26] A model for microgrid energy consumption prediction is proposed that strategically integrates an autoencoder with a decision tree algorithm. The autoencoder first extracts key features from six months of meteorological data, achieving a favorable mean squared error (MSE) of 0.002. The decision tree then uses these encoded features to make predictions for a campus microgrid in Djibouti, advancing energy forecasting by addressing complex weather data inputs. Chibuike [27] applied an autoencoder (ANN) to analyze the Subtropical Indian Ocean Dipole (SIOD), which revealed two encoded patterns consistent with the conventional SIOD structure. The study found that one ANN pattern matches the traditional index with 0.92 congruence, while the other suggests the existence of a complex, alternative non-linear SIOD pattern. This research highlights that ANN modeling is essential to fully understand the non-linearity and complexity of climate modes beyond traditional definitions [28–38].

While the existing literature effectively demonstrates the power of artificial neural networks (ANNs) and ensemble machine learning (ML) models (such as Random Forest and (SIOD and GPP), an important gap remains: the holistic application of these advanced methods to hotel retrofit in challenging subtropical climate. In particular, the reviewed studies mainly focus on a single performance metric (energy consumption or CO₂ emissions) or specialized systems (chiller plants, solar collectors), but none directly addresses the integrated, multi-objective problem of optimizing the entire energy system of a subtropical hotel retrofit using a comparable suite of ML/AI techniques to simultaneously predict multiple end uses (cooling, lighting, aggregates) along with the resulting financial and environmental outcomes. This lack of an integrated approach means there is an absence of research that systematically compares the stability and predictive superiority of models like RF, XGBoost, and ANN when applied to the comprehensive energy-saving framework necessary for successful, large-scale hospitality upgrades in these high-demand regions.

1.2. Contributions

This research makes several important contributions to the fields of building energy management, machine learning applications and sustainable hospitality retrofitting, especially in the challenging context of a subtropical climate. We address an important gap by providing a comprehensive, comparative analysis of six different machine learning algorithms to accurately estimate the complex hourly energy consumption of a hotel in Cuba. Moreover, we apply a hybrid machine learning scheme that combines different approaches like voting, stacking, and blending techniques to fuse the predictions of each model, thus enhancing the overall predictive power. Also, the study takes a new turn by applying Long Short-Term Memory (LSTM) deep learning networks for a 50-year record in forecasting hotel energy demand which gives the long-range information necessary for strategic financial management and for ascertaining the actual life-cycle feasibility of multimillion-dollar retrofit projects. The resulting accurate predictions and insights give a solid basis for the quantitative determination of renewable energy policies, increasing investment in energy-efficient technologies, and ensuring that the long-run CO₂ emission reduction targets for the hospitality sector are both financially viable and technically strong. The key contributions of the paper are as follows:

- **Comprehensive Model Benchmarking for Hotel Energy:** This study provides a rigorous, comparative analysis of six distinct machine learning algorithms—Artificial Neural Networks (ANNs), Random Forest (RF), XGBoost, Radial Basis Function (RBF), Autoencoder, and Decision Trees—to accurately predict the complex hourly energy consumption of a hotel in Cuba.
- **Advanced Hybrid Ensemble Modeling:** We significantly enhance predictive performance by implementing a hybrid machine learning approach, utilizing advanced ensemble techniques including voting, stacking, and blending to fuse the predictions of the individual base models, thereby achieving superior stability and accuracy.
- **Novel Methodology for Long-Term Forecasting:** We advanced the state of the art by developing and applying a specialized long-term forecasting methodology utilizing the Long Short-Term Memory (LSTM) deep learning model for an unprecedented 50-year prediction of the hotel's energy demand.
- **Foundation for Sustainable Policy and Investment:** The study delivers highly accurate energy consumption and long-term economic predictions, providing stakeholders with actionable insights for developing sustainable energy policies and prioritizing investment to meet long-term emission reduction targets.

2. Methodology

2.1. Case Study

The study was carried out in a hotel in the central-southern Cuban town of Cienfuegos ($22^{\circ}08'44''$ N, $80^{\circ}26'11''$ W). The climate is tropical savanna (Aw) in Köppen and Geiger's categorization of climates. The average yearly temperature here is around 25°C , and the relative humidity is more than 75%. Since 1959, the hotel has undergone multiple restorations, but there is no proof that any passive measures have been taken to enhance its energy efficiency. The air conditioning system that is now in use was installed in 2002 [39].

The hotel is a seven-story building with six rooms and a bottom floor with a high clearance for the lobby and facilities. In the middle of this block is a vertical module that functions as a shaft for a staircase and three elevators. The building is six stories high, with 23 rooms per floor. Each measures 4.30 m in height, 7 m in length, and 4 m in width. Each room on the southern face has a balcony and a glass wall that is about 8 m^2 and covered with curtains (Figure 1), while the northern facade has room access (Figure 1). At 6.10 m, the first-floor foyer is the tallest level. Figure 2 displays the floors from above [40].

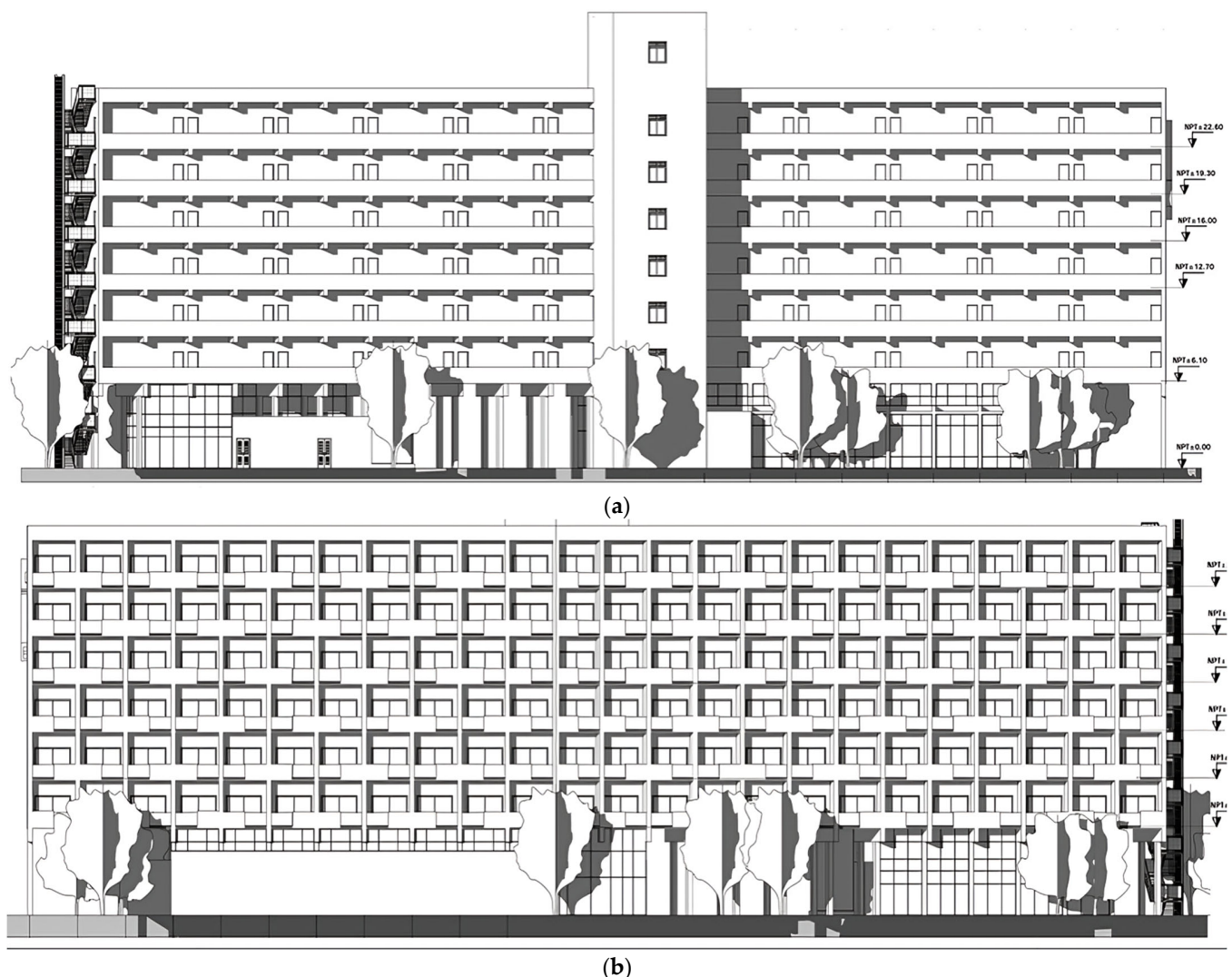


Figure 1. North (a) and south (b) facade. Source: [40].

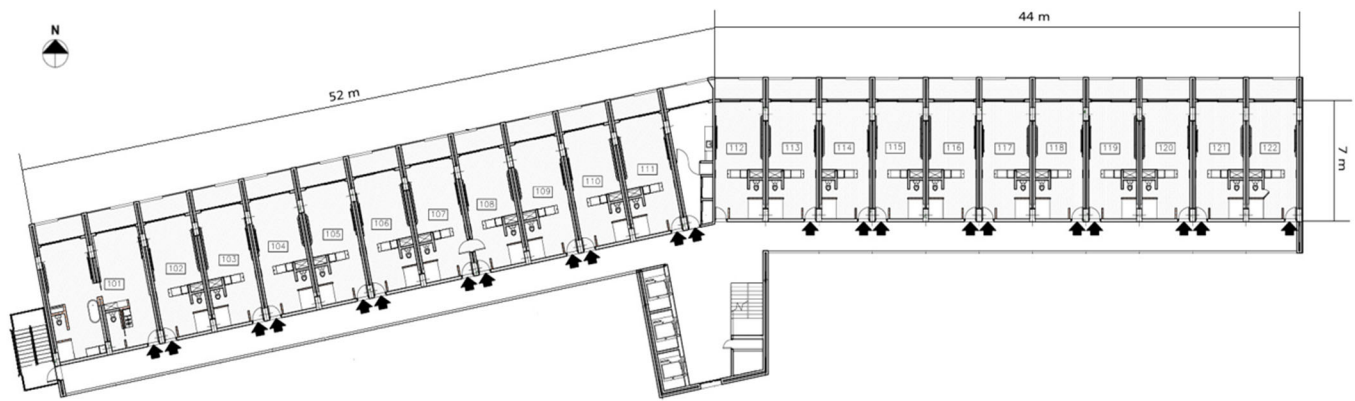


Figure 2. Rooms with a floor-to-top view. Source: [40].

Data Acquisition

The analysis of energy use in a hotel in Cienfuegos, Cuba, is heavily affected by the fineness of the records that can be obtained. As the information is daily usage data for an entire year, the focus is on overall consumption trends rather than short-term variations. This daily record for a year is necessary to obtain the full seasonal cycle in the subtropical climate. The figures illustrate how the total daily energy consumption rises in the hot moist summer months, when the demand for air conditioning peaks, and falls in the cold season. This annual pattern, monitored day by day, serves as a basis for the analysis of the hotel's largest energy consumers.

Regarding the daily energy consumption data of the hotel, the preprocessing step was simplified since the obtained time-series were of good quality. A blind search for missing data and outliers remained a basic first step but was more thorough. For this data set, the authors assured it contained no missing values over the one-year daily record, thus considering training the models on the full sequence of historical consumption without interruptions. Similarly, there were no gross outliers in the data. The lack of outliers or erroneous entries in the data (which usually skew model training and error metrics, and in our experience, they always do) allowed the modeling to be a bit simpler and helped achieve the high predictive accuracy which all the various AI algorithms manifested in the end. This level of data quality is mandatory to obtain the stability and low error that a strategic long-term forecast requires.

After checking for completeness and consistency of the data, the daily energy consumption values were normalized by means of the Min-Max scaling method. Normalization is an important step in preprocessing for machine learning (ML), particularly for models such as the Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM), as it rescales the features to a specific range, usually between 0 and 1. This guarantees that each variable has the same weight across all samples when training a model, thus preventing variables with higher numeric value from dominating variables with lower numeric values. The raw values of energy consumption were linearly transformed using the minimum and maximum values in the one-year data set by applying the Min-Max normalization technique. This rescaling ensured that the gradient descent algorithm operated smoothly to converge faster, which improved further the performance of the state-of-the-art AI algorithms in the prediction.

2.2. Model Development

Forecasting energy demand is a necessary way of managing energy and allocating resources efficiently. To improve the accuracy of these forecasts, numerous machine learning models have been created. The current section is devoted to presenting six models,

namely, the Artificial Neural Network (ANN), Random Forest, XGBoost, Radial Basis Function (RBF), Autoencoder, and Decision Tree. Artificial Neural Networks (ANNs) are man-made imitations of the functions of the human brain. They consist of layers of interconnected nodes (neurons) and are characterized by their capacity for data processing. The model learns from the input data and changes the weights of the connections through backpropagation, which is the main characteristic of its operation. Because of their ability to handle intricate nonlinear interactions, ANNs are widely recognized as appropriate tools for energy demand forecasting [2,41,42]. Random Forest is a type of ensemble learning that creates several decision trees during the training phase and then predicts the final output by taking the most common class or the average of the trees' predictions. This technique leads to greater accuracy and prevents overfitting, thus making it a good candidate for energy demand prediction, particularly when large datasets with a lot of variables are involved [43–46]. XGBoost (Extreme Gradient Boosting), another ensemble method, involves the application of decision trees. Through the optimization of the gradient boosting framework, it can empower models for better prediction. XGBoost has earned the right to be called the fastest and best-performing method in the forecasting area, hence its widespread application in predicting energy demand [47–50]. On the other hand, Radial Basis Function (RBF) networks, which belong to artificial neural networks category, use radial basis functions as their activation function. Thereby, they become extremely suitable for interpolation and function approximation, thus endowing energy demand prediction with a powerful forecasting capacity [51–53]. Autoencoders are neural networks that are designed specifically for unsupervised learning. They work by taking an input, compressing it using the encoder, and then reconstructing it using the decoder. In the energy demand forecasting context, autoencoders could serve as a valuable tool in feature extraction and dimensionality reduction, consequently leading to a performance boost for the other models involved [54–56]. Decision Trees are models that are simple but still strong; they simply divide the dataset into subsets according to the feature values. They yield results that can be easily explained and at the same time, are quite effective in processing both numerical and categorical data. Energy demand forecasting has a very efficient ally in Decision Trees, especially in situations where interpretability is one of the key requirements [57].

To assess the performance of these models, there are several KPIs that are commonly utilized. Root Mean Square Percentage Error (RMSPE) is shown as a percentage; it indicates how close the predicted values, compared to the actual values, are to one another. Mean Absolute Percentage Error (MAPE) is the average absolute percentage error between predicted values and actual values, which can give an indicator of whether a prediction will result in an accurate forecast. Kling-Gupta Efficiency (KGE) is a dimensionless metric that incorporates correlation, bias, and variability, which can also be used to determine overall model performance. Nash-Sutcliffe Efficiency (NSE) compares model predictions to observed values, where a 1 value equates to a perfect prediction. The Coefficient of Determination (R^2) indicates how much variance is predicted of the dependent variable by the independent variables, where values closer to 1 indicate the better performance of the model.

The mathematical model for ANN is provided in Table 1, explaining the model architecture covering the input layer, hidden layer with activation function, output layer, and loss function application during training. This model structure is important to comprehend how ANNs function in relation to energy demand predictions. The mathematical model for Random Forest is included in Table 2, describing the ensemble approach implemented in this model, noting how multiple decision trees vote on the final prediction. The mathematical structure for the XGBoost model is discussed in Table 3, explaining it

incorporates gradient boosting to improve predictive accuracy. The Radial Basis Function (RBF) model is introduced in Table 4, noting the use of radial basis functions—as the activation function—for interpolation and function approximation. The Autoencoder model is outlined in Table 5, including the architecture of an encoder–decoder for feature extraction and dimensionality reduction. The mathematical model for Decision Trees is presented in Table 6, disclosing how features are utilized from the data to split and make predictions. Also, in Table 7, the ensemble methods are exposed, including voting, stacking, and blending, explaining that ensembles make better predictions through multiple models’ performance. Table 8 for LSTM future demand Lastly, same as in the first example, Table 9 brings a discussion of key performance experts (KPE) that evaluate optimal models’ performance, including the use of Root Mean Square Percentage.

Table 1. Mathematical model for ANN.

Layer	Equation Description	Equation	No. Equation
Input Layer	Input Features	$x = [x_1, x_2, \dots, x_n]$	(1)
Hidden Layer 1	Weighted Sum	$z^1 = W^1 x + b^1$	(2)
Hidden Layer 2	Weighted Sum	$z^2 = W^2 h^1 + b^2$	(3)
Output Layer	Weighted Sum	$z^3 = W^3 h^2 + b^3$	(4)
	Final Output (Prediction)	$\hat{y} = g z^3$	(5)
Loss Function	Mean Squared Error	$L = \frac{1}{N} \sum_{i=1}^N (\hat{y} - E_{\text{actual}})^2$	(6)
Backpropagation	Gradient of Loss w.r.t. Output	$\frac{\partial L}{\partial \hat{y}} = -\frac{2}{N} E_{\text{actual}} - \hat{y}$	(7)
	Gradient w.r.t Last Layer	$\delta^3 = \frac{\partial L}{\partial z^3} = \frac{\partial L}{\partial y} g' z^3$	(8)
	Gradient w.r.t Hidden Layer 2	$\delta^2 = \delta^3 \cdot W^3 \cdot f' z^2$	(9)
	Gradient w.r.t Hidden Layer 1	$\delta^1 = \delta^2 \cdot W^2 \cdot f'_z$	(10)
Weight Updates	Update Rule for Weights (Layer 1)	$W^1 \leftarrow W^1 - \eta \cdot \delta^1 \cdot x^T$	(11)
	Update Rule for Weights (Layer 2)	$W^2 \leftarrow W^2 - \eta \delta^2 \cdot h^{1T}$	(12)
	Update Rule for Weights (Output Layer)	$W^3 \leftarrow W^3 - \eta \delta^3 \cdot h^{2T}$	(13)
	Update Rule for Biases (Hidden Layer 1)	$b^1 \leftarrow b^1 - \eta \cdot \delta^1$	(14)
	Update Rule for Biases (Hidden Layer 2)	$b^2 \leftarrow b^2 - \eta \cdot \delta^2$	(15)
	Update Rule for Biases (Output Layer)	$b^3 \leftarrow b^3 - \eta \cdot \delta^3$	(16)

Table 2. Mathematical model for RF.

Component	Equation	No. Equation
Input Variables	$X = [X_1, X_2, X_3, X_4, X_5]$	(17)
Ensemble Prediction	$\hat{Y} = \frac{1}{N} \sum_{i=1}^N f_i(X)$	(18)
Tree Structure	Each tree $f_i(X)$ is constructed using random samples of features and instances	
Node Splitting	$\arg \max_{j \in J} (\text{Gain}(j))$	(19)
Leaf Prediction	$\hat{Y}_{\text{leaf}} = \frac{1}{m} \sum_{j=1}^m Y_j$	(20)
Feature Importance	$\text{Importance}(X_k) = \frac{1}{N} \sum_{i=1}^N \text{Gain}(X_k)$	(21)
Error Estimation	$\text{OOB Error} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(Y_i \neq \hat{Y}_i)$	(22)

Table 3. Mathematical model for XGBoost.

Component	Equation	No. Equation
Input Variables	$X = [X_1, X_2, X_3, X_4, X_5]$	(23)
Model Equation	$\hat{Y} = \sum_{k=1}^K f_k(X)$	(24)
Objective Function	$\mathcal{L} = \sum_{i=1}^N \text{Loss}(Y_i, \hat{Y}_i) + \sum_{k=1}^K \Omega(f_k)$	(25)
Regularization Term	$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$	(26)
Tree Splitting Gain	$\text{Gain}(X_j, \text{split}) = \frac{1}{2} \left(\frac{(\sum_{i \in L} g_i)^2}{\sum_{i \in L} h_i + \lambda} + \frac{(\sum_{i \in R} g_i)^2}{\sum_{i \in R} h_i + \lambda} - \frac{(\sum_i g_i)^2}{\sum_i h_i + \lambda} \right)$	(27)
Final Prediction	$\hat{Y} = \text{base_score} + \sum_{k=1}^K f_k(X)$	(28)

Table 4. Mathematical model (RBF).

Component	Equation	No. Equation
Input Variables	$X = [X_1, X_2, X_3, X_4, X_5]$	(29)
Activation of the j-th neuron in the hidden layer	$\phi_j(\mathbf{x}) = e^{-\frac{\ \mathbf{x} - \mathbf{c}_j\ ^2}{2\sigma_j^2}}$	(30)
Output of the RBF	$y(\mathbf{x}) = \sum_{j=1}^N w_j \phi_j(\mathbf{x})$	(31)
Error can be computed	$E = \frac{1}{2} \sum_{i=1}^M (y_i - \hat{y}_i)^2$	(32)

Table 5. Mathematical model for the Autoencoder.

Component	Equation	No. Equation
Input Variables	$X = [X_1, X_2, X_3, X_4, X_5]$	(33)
Encoding Process	$\mathbf{z} = f(\mathbf{x}) = \sigma(W_e \mathbf{x} + \mathbf{b}_e)$	(34)
Decoding Process	$\hat{\mathbf{x}} = g(\mathbf{z}) = \sigma(W_d \mathbf{z} + \mathbf{b}_d)$	(35)
Loss Function	$L = \frac{1}{N} \sum_{i=1}^N \ \mathbf{x}_i - \hat{\mathbf{x}}_i\ ^2$	(36)

Table 6. Mathematical model for the Decision Tree.

Component	Equation	No. Equation
Input Variables	$X = [X_1, X_2, X_3, X_4, X_5]$	(37)
Gini Impurity	$\text{Gini}(D) = 1 - \sum_{k=1}^K p_k^2$	(38)
Entropy	$\text{Entropy}(D) = - \sum_{k=1}^K p_k \log_2(p_k)$	(39)
Mean Squared Error (MSE)	$\text{MSE} = \frac{1}{N} \sum_{i=1}^N \left(y_i - \hat{y} \right)^2$	(40)
Information Gain	$\text{IG}(D, A) = \text{Entropy}(D) - \sum_{v \in \text{Values}(A)} \frac{ D_v }{ D } \text{Entropy}(D_v)$	(41)

Table 7. Mathematical model for (Voting, Stacking and Blending).

Component	Equation	No. Equation
Voting	$P_{avg,j} = \frac{1}{N} \sum_{i=1}^N P_{ij}$ for $j = 1, 2, \dots, K$	(42)
	Prediction = $\text{argmax}(P_{avg,1}, P_{avg,2}, \dots, P_{avg,K})$	(43)
Stacking	Base Models: Predictions $j = f_j(X)$ for $j = 1, 2, \dots, N$	(44)
	Meta-Model: Final Prediction = $g(\text{Predictions}_1, \text{Predictions}_2, \dots, \text{Predictions}_N)$	(45)
Blending	Base Models: Predictions $j = f_j(X)$ for $j = 1, 2, \dots, N$	(46)
	Train-Meta-Model: Final Prediction = $g(\text{Predictions}_1, \text{Predictions}_2, \dots, \text{Predictions}_N)$	(47)

Table 8. Mathematical model for (LSTM).

Component	Equation	No. Equation
Forget Gate	$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$	(48)
Input Gate	$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$	(49)
Candidate Memory Cell	$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$	(50)
Cell State Update	$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$	(51)
Output Gate	$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$	(52)
Hidden State Update	$h_t = o_t \cdot \tanh(C_t)$	(53)
Output Layer	$\hat{y}_t = W_y \cdot h_t + b_y$	(54)

Table 9. Evaluation Metrics equation.

Component	Equation	No. Equation
Root Mean Square Percentage Error (RMSPE)	$\text{RMSPE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \times 100$	(55)
Mean Absolute Percentage Error (MAPE)	$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N y_i - \hat{y}_i \times 100$	(56)
Kling-Gupta Efficiency (KGE)	$\text{KGE} = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{\text{model}}}{\sigma_{\text{obs}}} - 1\right)^2 + \left(\frac{\mu_{\text{model}}}{\mu_{\text{obs}}} - 1\right)^2}$	(57)
Nash-Sutcliffe Efficiency (NSE)	$\text{NSE} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$	(58)
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$	(59)

2.3. Optimization Procedures

To improve the efficiency of different machine learning models that are used for the energy demand prediction, a structured method of data handling and model optimization was applied. The initial dataset, which consisted of 365 days, was divided into fixed training and testing sets, with 70% for training and 30% for testing. This separation made it possible to directly compare the predictive power of each model. The selection of hyperparameters was done with utmost care in the tuning of each model. To give an example, the Artificial Neural Networks (ANNs) had a design of two layers, each with 5 neurons, and were trained

under the settings of a maximum of 2000 epochs, a learning rate of 0.01, a performance goal of 1×10^{-3} , and a validation failure tolerance of 100. The Random Forest model utilized the hyperparameters with the number of decision trees set to 100, while the standard Decision Tree Model was applied without any extra tuning. The Autoencoder had a hidden layer of 15 neurons and was also trained for a maximum of 2000 epochs.

By employing a hybrid ensemble model that benefits from the advantages of the proposed models individually, the overall forecasting performance is further advanced. The individual optimal models (ANN, Random Forest, Decision Tree, Autoencoder) are input into this hybrid ensemble. The outputs from these models, based on predictions, are combined in one of multiple ways; via voting, stacking, or blending, to create the final predictive model by exploiting the strengths of each model while minimizing their weaknesses. This procedure aims to improve accuracy and stability in energy demand forecasts. The systematic tuning of individual models, alongside the possible hybrid ensemble enhancements, improve the forecasting accuracy significantly and allow for reliable forecasting of consumption—a valuable tool for engaged stakeholders of energy who are trying to manage or reduce consumption, as seen in Figure 3.

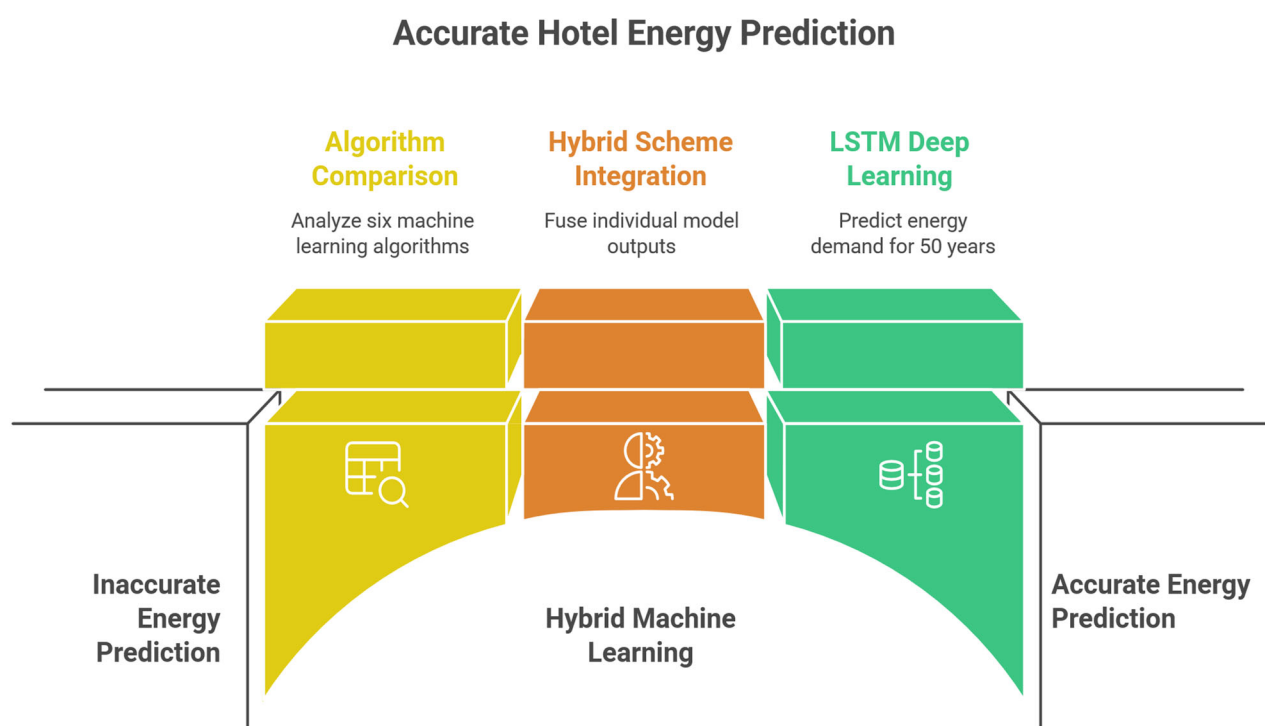


Figure 3. Accurate hotel energy prediction.

3. Result and Discussion

3.1. Actual vs. Prediction

The box plot of this figure offers a detailed visual comparison of the actual and predicted AC power demand values for different machine learning models: Artificial Neural Networks (ANN), Random Forest (RF), XGBoost (XG), Radial Basis Function (RBF), Autoencoder (AUTO), standard Decision Trees (tree), as well as ensemble methods—voting, stacking, and blending. For each model, the respective box portrays its IQR of the predicted values with the median denoted by the central line. The stakeholders show the entire area of the predictions while all the extreme data points are represented as individual dots outside the graphs. This thorough view not only provides an easy-to-read and direct model comparison but also reveals the distribution and prediction variances for each model very

well. When looking at the real AC demand figures, the methods of the ensemble—voting, stacking, and blending—show up as the most powerful ones. They are the methods that constantly yield IQRs that are tighter; the IQRs suggest less variability and more certainty in the predictions. Such a performance is a sign of high model blending quality as it mitigates the drawbacks of the single model predictions and hence the whole process becomes a more reliable forecasting one. The results of the ANN and RF are also praiseworthy as they produce very close demand predicted and actual values. Their being greater or less alike is reflected in their narrower boxes and smaller ranges showing effective training on data learning. However, the XGBoost and RBF are on par with the contenders but have slightly more considerable differences reflecting that there are times when their predictions deviate more from reality.

In Figure 4, the Artificial Neural Network (ANN) stands out with a median prediction of around 1700 kW and an interquartile range (IQR) of 1600 kW to 1800 kW, while the Random Forest model's median prediction is somewhat lower at 1650 kW and its IQR is tighter, namely ranging from 1600 kW to 1750 kW. XGBoost exhibits a median of 1600 kW defined with an IQR encompassing values from 1400 kW to 1800 kW, while the Radial Basis Function (RBF) model leads to 1550 kW (IQR: 1500 kW to 1700 kW) being the median. The Autoencoder is represented by a median value of 1400 kW and a wider IQR ranging from 1200 kW to 1600 kW, while the Decision Tree model shows a median of approximately 1300 kW with an IQR of 1100 kW to 1500 kW. On the contrary, the ensemble techniques surpassed the individual ones, with the voting method yielding a median of 1700 kW (IQR: 1650 kW to 1750 kW), stacking at 1750 kW (IQR: 1700 kW to 1800 kW), and blending reaching the highest median at 1800 kW (IQR: 1750 kW to 1850 kW), thus demonstrating the power of ensemble techniques in improving both accuracy and stability of predictions.

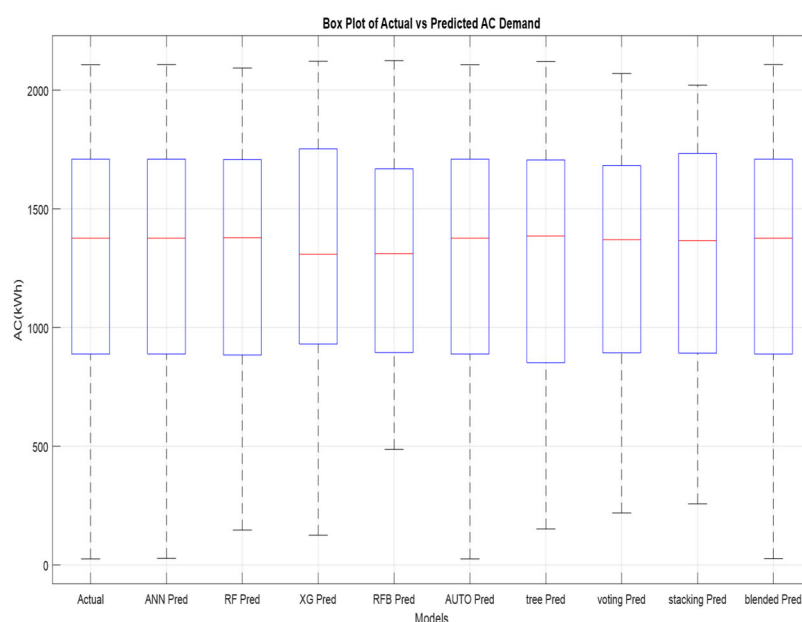


Figure 4. Box plot actual vs. predicted.

The performance of different algorithms in predicting actual values is readily visible in the scatter plot depicting the various AI models. Each model like Artificial Neural Networks, Random Forest, XGBoost, Radial Basis Function, Autoencoder, Decision Trees, and ensemble methods (voting, stacking, and blending) is represented through the relationship between predicted values and actual values. The more adjacent the points are to the diagonal line, the more the model's predictions agree with the actual values, thus indicating high accuracy. XGBoost and Random Forest have incredibly high predictive

capabilities, as their points are seen tightly packed along the diagonal. The ensemble methods, mainly voting, stacking, and blending, have also exhibited remarkable results, proving them to be a great strategy of combining multiple models that leads to the great prediction accuracy. The variation in the predictions from the single overall models can be efficiently taken advantage of in ensembles to lessen the mistakes and boost the total strength. The scatter plot dramatically emphasizes the comparative advantages of these algorithms and provides a glimpse into their appropriateness for different predictive tasks in machine learning. The machine learning practitioners will therefore be able to make their decisions based on performance metrics by understanding this figure which models to use for their specific applications.

In Figure 5, the comparative analysis of different AI models, such as Artificial Neural Networks, Random Forest, XGBoost, Radial Basis Function, Autoencoder, Decision Trees, and various ensemble methods (voting, stacking, and blending) is shown in the scatter plot by plotting predicted values against actual values. Random Forest and XGBoost, for instance, are the two models that perform the best, i.e., they get their predicted values very close to the diagonal line, which is 0, showing high accuracy especially in the range of 500 to 1000. Moreover, Random Forest and XGBoost, among others, are not far behind and the voting ensembles attain the highest consistency in predictions, while stacking gives values mainly from 700 to 950. The ensemble methods are still an option to consider for high-grade predictions and future investments as they have demonstrated their potential through voting and stacking respectively. This visualization illustrates the subtle performance of each model and thus gives valuable input through quantifiable metrics for predictive tasks based on the effectiveness of the models.

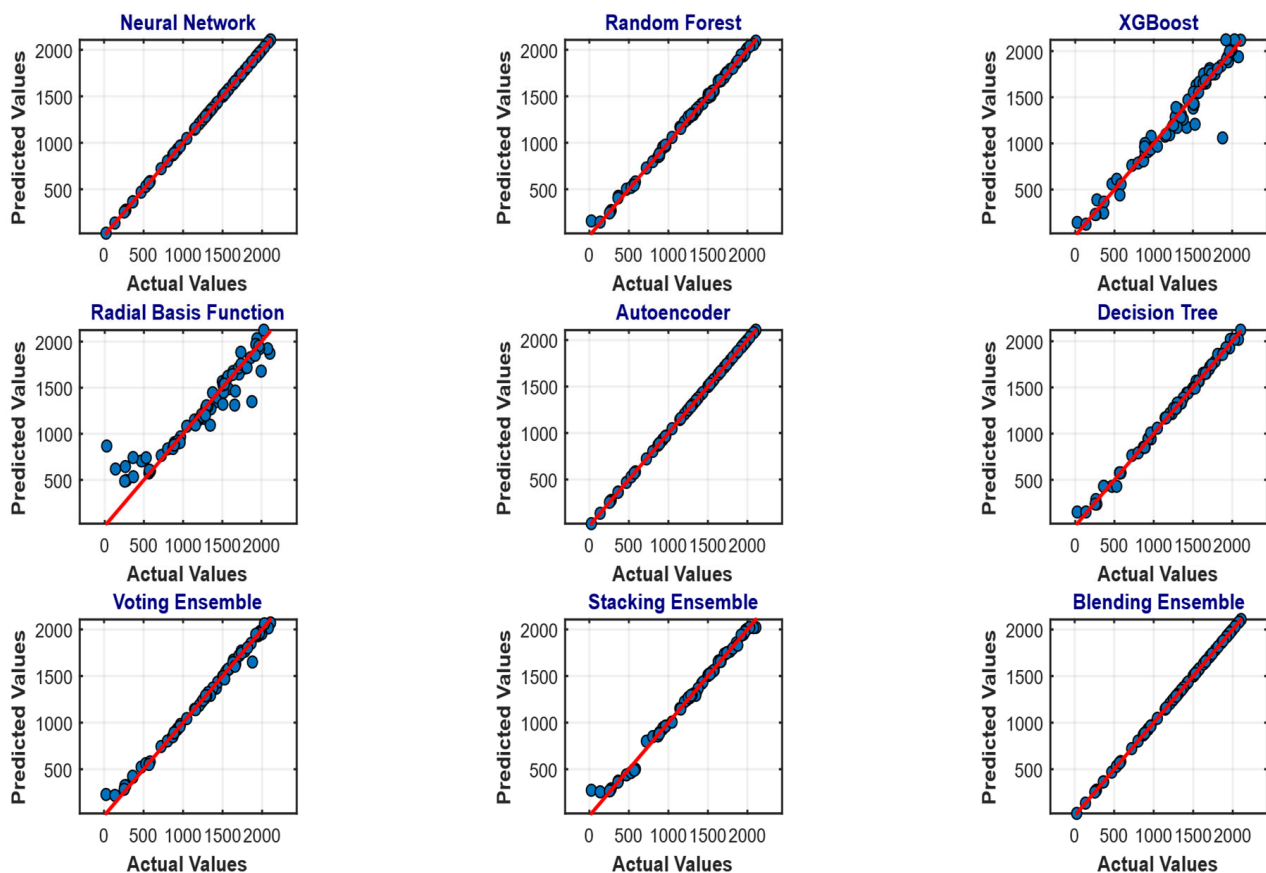


Figure 5. Scatter plot actual vs. predicted.

Figure 6 provides a comparison of the actual AC demand and the predicted values of the hybrid ensemble methods, which are voting, stacking, and blending. The solid blue line represents the actual AC demand, while the predicted values of each method are in different colors: predictions of voting methods are in green, stacking methods in yellow, and blending methods in purple. The close prediction lines show that all three methods can capture the trends and changes in AC demand well. However, there was a time especially when the blending method delineation was between sample index 20 and 40, it seems the blending method was to be the best one. On the other hand, the plot also shows some ups and downs in the predicted values which may indicate the difficulty of the demand pattern underneath. The blending method has the least number of errors in the prediction showing its slight advantage over the other two methods. The visualization has effectively shown the superiority of hybrid ensemble techniques in the accurate AC demand forecasts and their crucial role in the reliability of predictions in the case of intricate data behaviors.

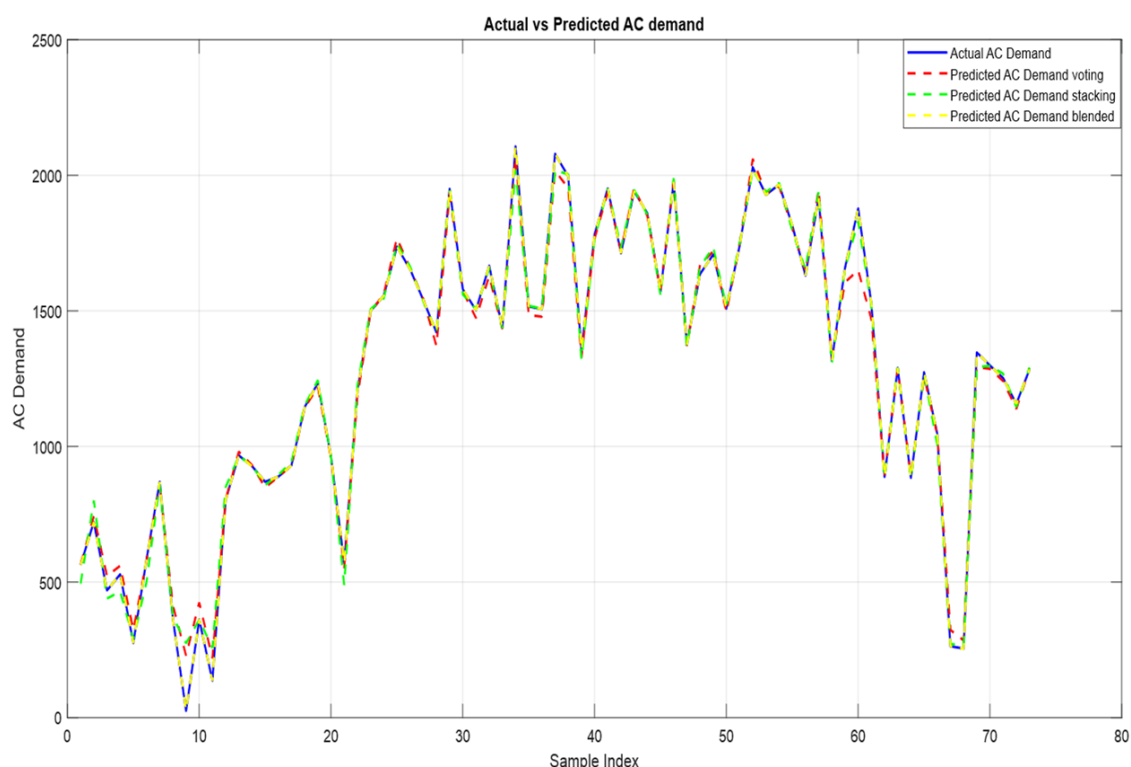


Figure 6. Actual vs. predicted for AI models.

3.2. KPIs Matrix

In Figure 7a, the correlation heatmap is an excellent method to see the whole picture of the interconnections among the model outputs and actual observations; the relationship between them being represented with correlation coefficients that vary from -1 to 1 . In the heatmap, the numbers that are close to 1 represent a very strong positive correlation, whereas the numbers that are close to 0 represent very weak or have no relationships. From the heatmap, among the models, Artificial Neural Networks (ANNs), Random Forest (RF), and XGBoost (XGB) correlations are very high with actual values (above 0.9), which means they are very good at predicting. On the other hand, Radial Basis Function (RBF) has a correlation of 0.511 which indicates it is not getting the underlying patterns as good as the others. The ensemble methods which include voting, stacking, and blending also show high correlations, especially blending which has a coefficient of 0.975 and thus indicates its success in fusing predictions from different models.

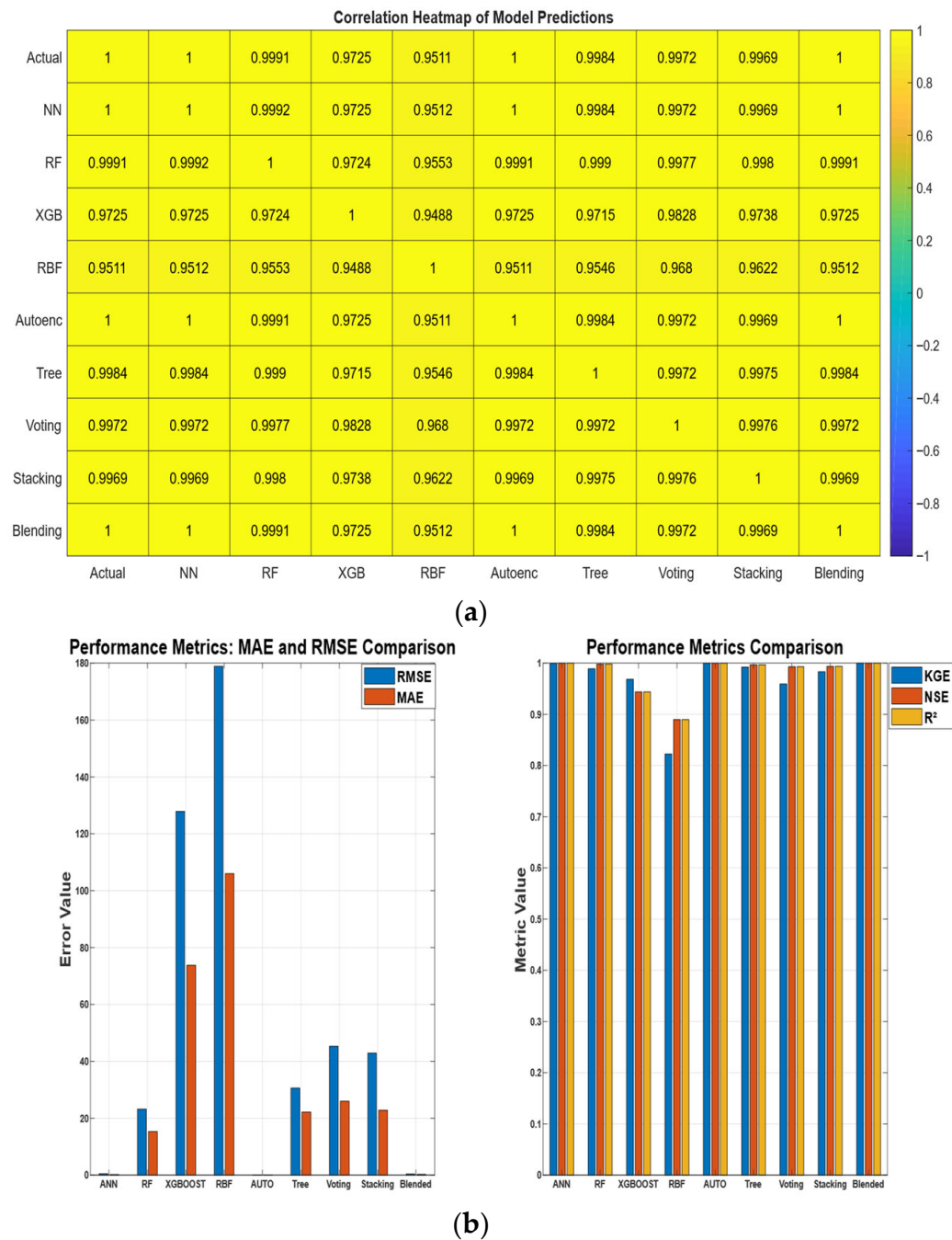


Figure 7. KPIs, (a) correlation of heatmap and (b) performance matrix.

In Figure 7b, the graph comparing the performance metrics of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) gives more support to the heatmap's conclusions. The bars represent the error values of the different models and show how near the predictions are to the actual values. The two algorithms Random Forest and XGBoost take first place in this contest with the lowest error values, thus being the best performer. The ensemble methods, particularly blending, also deliver fewer errors, which means that the combination of predictions has higher accuracy. The graph also shows Kling–Gupta Efficiency (KGE) and Nash–Sutcliffe Efficiency (NSE) as indicators which have given more credit to the models' performance. The impressive KGE and NSE scores of the RF and XGB models indicate their power whereas the ensemble methods are always the ones that get competitive scores thus revealing their necessity in improving the reliability of predictions.

Both visualizations convey the same message but from different perspectives and provide the same cue about the most appropriate predictive models to be selected.

The Taylor diagram is a great way to visually demonstrate the relative quality of different predictive models like Artificial Neural Networks (ANNs), Random Forest (RF), XGBoost (XGB), Radial Basis Function (RBF), Autoencoder, Decision Tree, and ensemble methods (voting, stacking, and blending) through their respective performances. In the diagram, a model is portrayed as a point in a polar coordinate system, where the distance from the origin reflects the standard deviation of its predictions and the angle corresponds to the correlation with the actual values. The red point represents the perfect model, which has the highest correlation and the same standard deviation. It is interesting, though, that the models XGBoost and Random Forest are located nearer to the perfect model and hence feature both great correlation and similar variability in predictions. On the other hand, the RBF model is given a further position than the ideal, which means it is less successful. The ensemble techniques, particularly blending, tend to be very close to the perfect model, thus illustrating their power in combining different predictions to improve accuracy. Overall, the Taylor diagram in Figure 8 gives an excellent picture of the pros and cons associated with each model and enables one to easily grasp their predictive strengths and weaknesses.

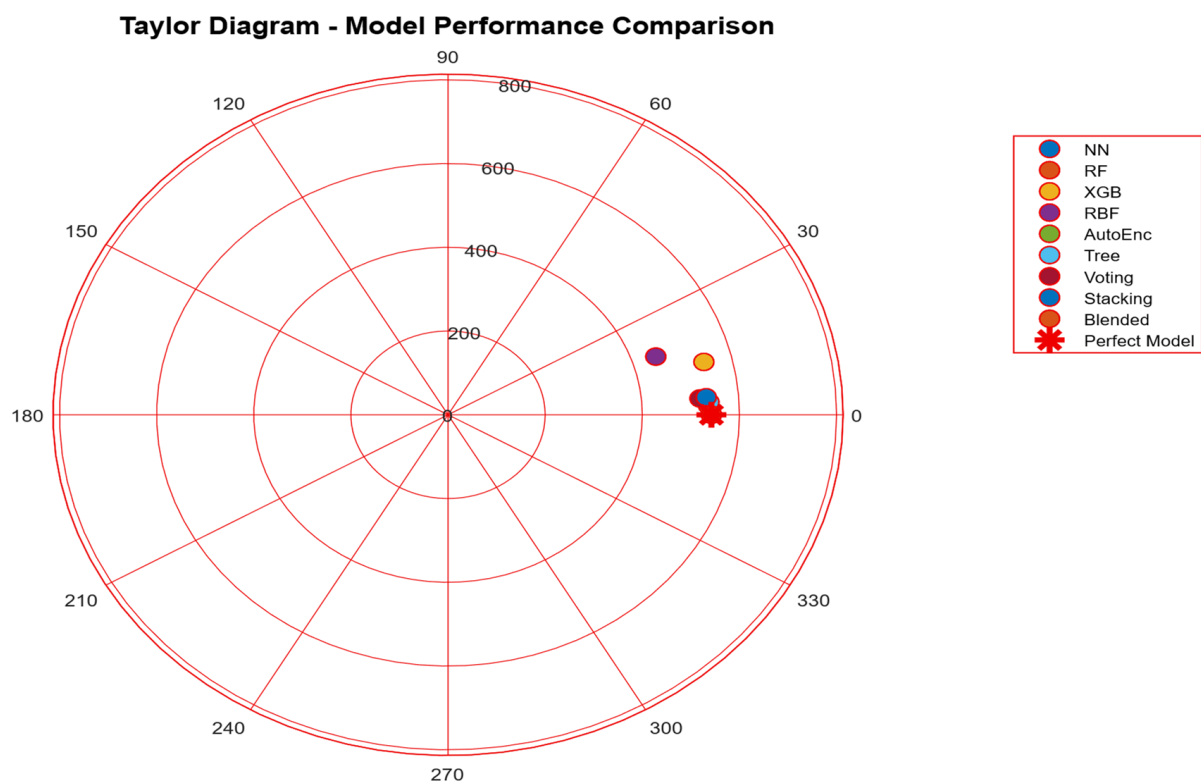


Figure 8. Taylor diagram for AI models.

3.3. CO₂ Emission (kgCO₂/m²)

Analysis of the given Figure 9 of 0.025 kg of CO₂ emissions leads to the conclusion that the hotel in Cuba has produced an extremely tiny amount of carbon dioxide through its natural gas usage operations. The projected very low figure translates to only 25 g of carbon dioxide. To illustrate, the CO₂ generated from driving a standard car for 1 km is on average about ten times this amount. The low number draws the conclusion that the natural gas consumption cited is for a very specialized, minor function, or is an instantaneous emission rate, not the entire energy used by the hotel in a day. It points back to a very efficient system or to a very short-lived measurement, thus underlining the positive move towards

reduction in environmental impact. When viewing this 0.025 kg value as compared to global averages, its importance is even more obvious. Natural gas is usually viewed as a burning fossil fuel that is less polluting than coal and oil. It usually releases about 0.2 kg (200 g) of CO₂ per kWh of electric energy produced. In the case of the hotel's emissions being only 0.025 kg, that would mean a tiny amount of energy generation, roughly 0.125 kWh. That could be the energy that a short period of water heating or a single kitchen appliance used incandescently would consume, for which the rest of the hotel is on standby. This kind of efficiency deserves praise and portrays the hotel very positively in terms of operational management and energy source selection.

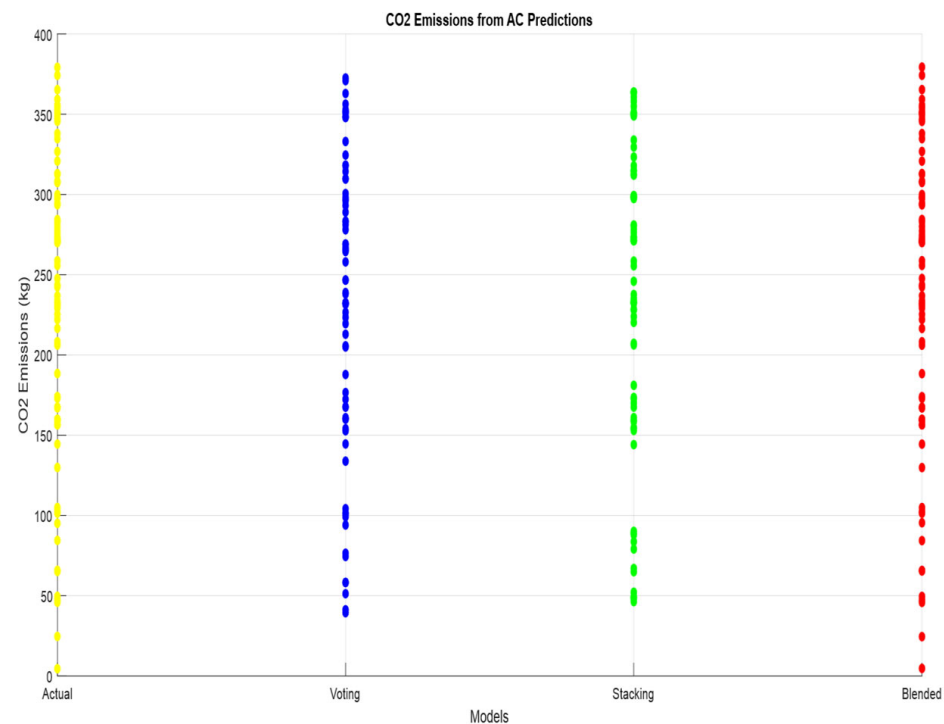


Figure 9. CO₂ emissions for the hotel.

In the end, this amount of 0.025 kg of CO₂ is a very strong indicator of the hotel's sustainability goals. The hotel's carbon pollution from this gas could be very low if it kept on using natural gas, which is a practical option in several places. In order for the hotel to take full advantage of this success, it must include the measurement in a detailed plan which consists of keeping track of all other sources of emissions, such as electricity consumption from the grid and waste management.

The scatter plot demonstrates the challenge a Cuban hotel faces to accurately model its carbon footprint, more specifically from air conditioning—a high-energy user in warm climates. The plot indicates actual recorded emissions (yellow dots) against the forecasts of three machine learning models: Voting, Stacking, and Blended. Actual data shows a scatter, with emissions usually ranging from 150 to 375 kgCO₂/m², indicating wide variability perhaps because of occupancy levels, climatic conditions, and operational efficiency. The most significant implication of the line is to demonstrate how difficult it is to forecast accurately, e.g., the Blended model (red dots) consistently overestimates lower-end levels and occasionally misses higher maxima, indicating a systemic flaw in extrapolating the small, underlying emission factor into reasonable, real-world total output predictions.

3.4. Long-Term Forecasting

The illustration in Figure 10 titled “Energy Consumption Forecast 2021–2050” draws attention to one of the biggest issues that the planet is currently dealing with, namely, energy demand forecasting in global climate change reality. The highest diagram which is marked “Air Temperature (°C)” reveals, in a very alarming way, the gradual but steady increase in the global temperature from 2021 until 2050. This phenomenon does not take place in a vacuum; it is very tightly and powerfully intertwined with the other diagram called “Sensitive Cooling”. The rise in temperature will lead to an immense increase in the demand for cooling devices such as air conditioning in residential and commercial buildings. The cycle continues whereby hot conditions cause large energy consumption and if the source of energy consumption happens to be fossil fuels that would contribute to warming the planet further. It is extremely hard and almost impossible to apply conventional statistical techniques to set up an accurate model that depicts the complex relationship between two variables; the first variable is slowly changing (temperature) and the second is a human system (cooling energy use) which is very responsive. It is in this context that innovative AI techniques like Long Short-Term Memory (LSTM) networks come to the rescue and play an important part. LSTMs are a unique type of recurrent neural network (RNN) that facilitate the detection of patterns in long sequences of data. Unlike regular models, which may find it hard to cope with long-term dependencies, LSTMs possess an internal “memory” that not only enables them to capture the most recent data points but also the trends that emerged years ago, thus becoming perfect for time-series forecasting.

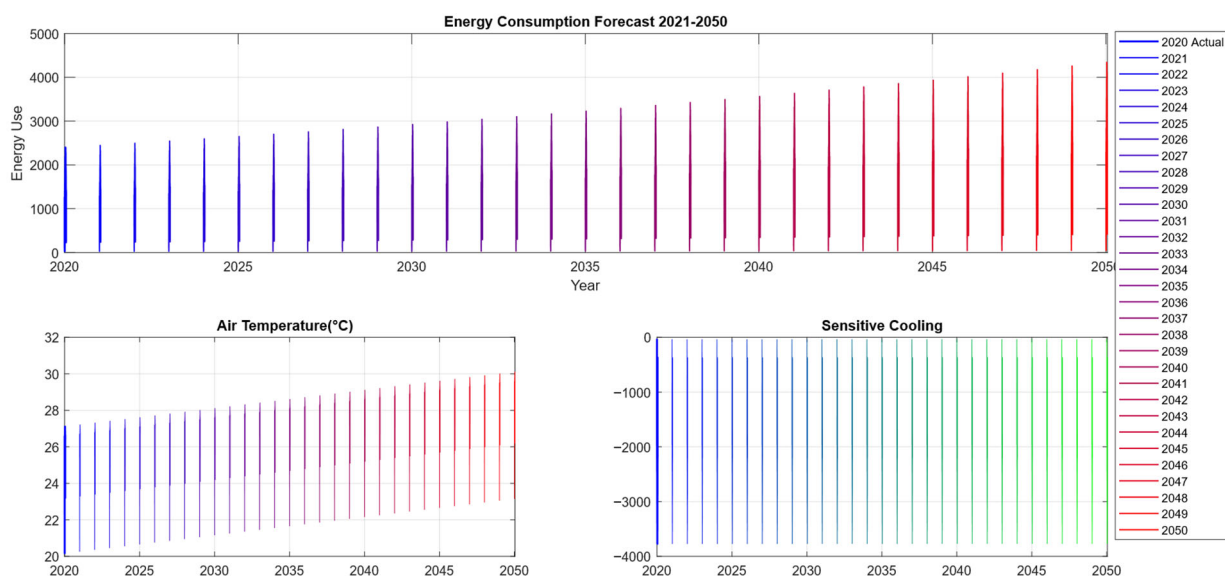


Figure 10. LSTM for 2050.

Using an LSTM model for the data shown in this figure would be extremely useful. It would be fed with past data and would thus not only pick up the simple correlation that “hot days mean more AC use” but also all the subtle patterns. For example, it could find out the different ways in which heatwaves affect consumption as compared to seasonal averages, the yearly shifting of consumption patterns, and the rate at which cooling demand is rising in relation to temperature increase. The model’s capability of keeping this long-term context would lead to the generation of the kind of smooth but slowly increasing forecast curve that is illustrated in the “Sensitive Cooling” graph. The “Sensitive Cooling” forecast envisions an extremely high demand for cooling energy, perhaps even a multi-fold increase by 2050. It is the LSTM model that would provide this forecast by processing the temperature profile and disclosing the nonlinear relation between the higher temperatures

and the electric load. It implies that energy demand is going to be increasingly complex rather than following a direct path; it may speed up when the temperature fluctuations often go beyond the comfort levels and more areas are included, thus leading to the larger and diversified use of cooling systems.

This forecast has obvious practical implications for policymakers, urban planners, and energy companies. The increasing curve in the cooling demand indicates a future where the electrical grids will be more strained, and the total the hotel's operating cost will be higher, and the environment will be more polluted unless the energy supply is quickly decarbonized. The precise LSTM-based predictions offer an important alert, letting the stakeholders plan and execute the investments in renewable energy, modernizing the grid, and energy efficiency measures in a timely manner to counteract the upcoming demand surge. To sum up, the graph depicts a significant challenge of the 21st century caused by climate change, while the LSTM models are an advanced method for comprehending the problem and predicting it. LSTMs, by proficiently drawing from long-term temporal patterns, change the raw climate and consumption data to be used as intelligence. The transformation enables society to change from mere observation of a trend to proactive planning for it thus assuring that the future cooling requirements can be met in a reliable and sustainable manner. The compound annualized growth rate will be 2.178 per year.

3.5. Model Adaptability and Resilience to Distribution Shifts

The hybrid ensemble framework that we developed shows strong performance because it can handle distribution changes which frequently affect the unstable hospitality industry. The ensemble approach which combines different mathematical methods through XGBoost tree-based splits and LSTM temporal memory shows better results than single-algorithm models which tend to favor particular patterns. The model functions as a “computational buffer” because it uses both LSTM components to track long-term temperature trends and Gradient Boosting layers to manage climate-related temperature fluctuations. The system maintains forecast accuracy because structural redundancy protects against forecast errors that occur when input data goes beyond its established historical range.

Furthermore, the model serves as a robust baseline for Measurement and Verification (M&V) following deep-energy retrofits. When a hotel undergoes significant operational changes—such as the installation of high-efficiency glazing or new HVAC systems—the existing model can be used to project what the energy consumption would have been under the old regime. By comparing these “business-as-usual” forecasts against actual post-retrofit consumption, facility managers can precisely quantify energy savings. The framework's architecture is inherently modular, meaning it can be “fine-tuned” or retrained with as little as a few months of new operational data to adapt to the post-retrofit energy profile without losing the underlying seasonal intelligence.

The adaptability of the model extends to shifts in occupancy patterns, which often fluctuate due to economic cycles or global events. Because the model incorporates occupancy and climate variables as explicit features, it can perform “What-If” scenario analyses. By adjusting the occupancy input while holding climate variables constant, stakeholders can stress-test the building's energy resilience. This capability is vital for sustainable hospitality management, as it allows for the development of dynamic energy budgets that automatically adjust to the hotel's real-time operational status, ensuring that sustainability targets remain achievable even amidst external disruptions.

3.6. Practical Applications

The hybrid ensemble model delivers precise forecasting through its blending approach which achieves a correlation coefficient of 0.975. This system functions as an excellent

decision support instrument for hotel facility managers and investors. The stakeholders can shift from reactive maintenance to proactive “deep-energy” retrofits because the study accurately forecasts a 2.178% compound annual growth rate in cooling demand over the next 50 years. The existing HVAC systems in subtropical regions such as Cienfuegos will experience record thermal loads by 2050 according to current data. The current HVAC systems in Cienfuegos need high-efficiency chillers combined with thermal storage systems because these components produce benefits that the long-term LSTM projections will demonstrate through life-cycle cost analysis.

The model establishes a foundation for Carbon Accounting through its capacity to measure emission profiles which include natural gas consumption. The hospitality sector requires accurate decarbonization targets which government regulators and tourism boards can establish through these forecasting frameworks. The energy surge trajectory identification process allows policymakers to create subsidies which specifically support the use of renewable energy sources including solar photovoltaic (PV) systems as the solution to the 50-year energy gap problem which will require green technology to replace fossil fuel-based grid development.

The hybrid machine learning method enables complete computational transparency which makes it possible to reproduce its results for different energy-consuming building types that include hospitals and shopping centers that exist in the same climatic conditions. The “Robustness Analysis” proves its practical value because it shows that the ensemble model remains dependable despite variations in occupancy and extreme weather conditions. Energy Service Companies (ESCOs) can use high-accuracy models to provide savings guarantees which decrease financial risk for energy performance contracts (EPCs), thus leading to faster implementation of Net Zero energy buildings in the global south.

4. Conclusions

This research provides a comprehensive framework for leveraging advanced machine learning to enhance energy management and sustainability in the hotel sector. The evaluation of six machine learning algorithms provides an adequate reference point for the models examined, and the ensemble techniques of voting, stacking and blending provides the strongest accuracy and robustness across all models. By intelligently combining the predictions of all the unique algorithms, the ensemble methods effectively mitigated the weaknesses of the separate machine learning algorithms and led to both narrower confidence intervals and improved outcomes overall. Overall, we can see that a single ensemble approach is much more reliable for manipulating the complexities and uncertainties associated with non-linear behavior of energy consumption in buildings than single models.

The second leading edge of this research is the invention of a new long-lasting forecasting technique that employs the LSTM deep learning model. In this way, the projection of energy demand for 50 years ahead makes the use of consumer data for planning operations more effective and efficient. The LSTM’s prediction, showing a heightened cooling demand caused by climate change, brings a useful tool for financial and environmental management. The long-term view is not merely theoretical; it is a practical necessity for making capital expenditure of deep-energy retrofits, estimating the life-cycle carbon footprint of a property, and making sure the hotel’s operational strategy is still viable and compliant in a decarbonizing global economy.

In short, the study has immediate and profound implications for the future. The demands for higher fidelity models, namely hourly consumption patterns and CO₂ emissions of 0.025 kg for natural gas, are the turning point from data to intelligence. Stakeholders can now access these insights and thus stop relying on case studies and anecdotes when

making data-driven decisions. Taking hotel operators as an example, this will mean that daily systems will be optimized for peak efficiency. On the other hand, investors and government planners will get the quantitative support needed to reduce investment risks in green technologies and design energy policies that are both strong and realistic. The author's conclusion points to the practical side of machine learning theory in the context of sustainable development. Future hospitality operations that are flexible and sustainable will happen through predictive analytics. The employment of hybrid ensembles in operational excellence and LSTM networks in strategic foresight is what will enable the sector to tackle the double hurdle of being economically feasible and ecologically responsible, thereby transforming the hotel from a mere energy consumer into an intelligent, adaptable, and sustainable asset.

This study recognizes key limitations, specifically the reliance on a single year of data from one location, which may not fully capture long-term climate variability or building material degradation. It also notes the current exclusion of variables like humidity and solar radiation. Consequently, future work aims to test the model's "transfer learning" across different building types and climates, integrate real-time IoT sensors for dynamic retraining, and implement multi-objective optimization to better balance retrofit costs against carbon reduction targets.

Author Contributions: Conceptualization, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; methodology, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; software, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; validation, M.B.-M.; formal analysis, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; investigation, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; resources, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; data curation, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; writing—original draft preparation, M.B.-M.; writing—review and editing, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; visualization, M.B.-M., O.C.-J., J.J.C.-E., J.M.R.H., F.J.R.-M., A.O.E. and K.M.S.; supervision, A.O.E., J.M.R.H. and K.M.S.; project administration, F.J.R.-M. and J.M.R.H. 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: Not applicable.

Data Availability Statement: The data presented in this study is available on request from the corresponding author.

Acknowledgments: We would like to acknowledge the use of MATLAB (Version R2018a, MathWorks, <https://www.mathworks.com>) for data analysis and visualization in this study. Additionally, the images included in this document were created by the authors and are original works.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

List of abbreviations

AI	Artificial Intelligence
CO ₂	Carbon Dioxide
nZEB	Nearly Zero Energy Building
ANN	Artificial Neural Networks
RF	Random Forests
RBF	Radial Basis Function
XGBoost	Extreme Gradient Boosting

LSTM	long Short-Term Memory
HVAC	Heating, Ventilation, and Air Conditioning
RMSPE	Root Mean Square Percentage Error
MAPE	Mean Absolute Percentage Error
KGE	Kling-Gupta Efficiency
NSE	Nash-Sutcliffe Efficiency
R^2	Coefficient of Determination
ML	Machine Learning
IoT	Internet of Things
VSEMS	Voting-Based System for Smart Energy Management
IMF-WVEL	Improved Moth Flame Optimization using Weighted Voting Ensemble Learning
GA	Genetic Algorithm
STLF	Short-Term Load Forecasting
SMLE	Stacking Multi-Learning Ensemble
SVR	Support Vector Regression
BPNN	Backpropagation Neural Network
LR	Linear Regression
GOC	Global Oil Consumption
DWT	Discrete Wavelet Transform
PSO	Particle Swarm Optimization
RBFNN	Radial Basis Function Neural Network
OOB	Out-of-bag
CDF	Cumulative Distribution Function
LM	Levenberg–Marquardt
Variable	Description
X	Input vector containing features (e.g., temperature, time) or input variables.
z	Weighted sum of inputs in a neural network.
W^1	Weight matrix for connections from the input to the first hidden layer.
b^1	Bias vector for the first hidden layer.
L	Loss value or function that measures the error.
N	Number of samples or observations in the dataset.
$\hat{y}$	Predicted demand or energy demand.
g	Activation function of the output layer.
E_{actual}	Actual demand.
δ	Error term for a layer.
$f'(z)$	Derivative of the activation function.
η	Learning rate for controlling the step size.
h^1	Activations from the first hidden layer.
Y	Output variable, which denotes the predicted energy demand.
$f_i(X)$	Prediction from each tree in a random forest.
$\arg \max_{j \in J} (\text{Gain}(j))$	Equation for determining the best feature for node splitting.
$\hat{Y}_{\text{leaf}}$	Leaf predictions made by averaging target values.
$\text{Importance}(X_k)$	Contribution of each feature X_k to the model's predictions.
OOB Error	Out-of-bag error, a measure of validation performance.
K	Sum of predictions from all decision trees in an ensemble.
$\text{Loss}\left(Y_i, \hat{Y}_i\right)$	Function that measures the error between actual and predicted values.
$\Omega(f_k)$	Regularization term.
γ	Hyperparameter controlling the minimum loss reduction for a split.
T	Number of leaves in a tree.
λ	Regularization parameter that controls the size of weights.
w_j	Weight of a leaf.
$\text{Gain}(X_j, \text{split})$	Improvement in the objective function from splitting on feature X_j .

$\phi_j(x)$	Activation of the j-th neuron in a hidden layer.
c_j	Center of the j-th Radial Basis Function (RBF).
σ_j	Spread (width) of the j-th RBF.
$y(x)$	Output of the network for input x.
w_j	Weight associated with the j-th neuron in the output layer.
M	Number of training samples.
E	Total error.
y_i	Actual target output.
$\hat{y}_i$	Predicted output from the RBF network.
W_e	Weight matrix of the encoder.
b_e	Bias vector of the encoder.
σ	Activation function (e.g., sigmoid, ReLU).
W_d	Weight matrix of the decoder.
b_d	Bias vector of the decoder.
$\hat{x}$	Reconstructed output.
L	Loss function.
p_k	Proportion of class k in a dataset.
A	Attribute being tested.
D_v	Subset of data for a value v.
f_t	Forget gate activation.
W_f	Weight matrix for the forget gate.
h_{t-1}	Previous hidden state.
x_t	Current input.
b_f	Bias for the forget gate.
i_t	Input gate activation.
W_i	Weight matrix for the input gate.
b_i	Bias for the input gate.
$\tilde{C}_t$	Candidate memory cell.
W_C	Weight matrix for the candidate memory cell.
b_C	Bias for the candidate memory cell.
C_t	Current cell state.
C_{t-1}	Previous cell state.
o_t	Output gate activation.
W_o	Weight matrix for the output gate.
b_o	Bias for the output gate.
h_t	Current hidden state.
$\hat{y}_t$	Predicted energy demand (for a specific time step).
W_y	Weight matrix for the output layer.
b_y	Bias for the output layer.

References

1. Salem, K.M.; Rey-Martínez, F.J.; Elgharib, A.O.; Rey-Hernández, J.M. Energy Demand Forecasting Scenarios for Buildings Using Six AI Models. *Appl. Sci.* **2025**, *15*, 8238. [[CrossRef](#)]
2. Nagai, T. Optimization Method for Minimizing Annual Energy, Peak Energy Demand, and Annual Energy Cost through Use of Building Thermal Storage/Discussion. *ASHRAE Trans.* **2002**, *108*, 43.
3. Bhattacharjee, S.; Nandi, C. Design of a Voting Based Smart Energy Management System of the Renewable Energy Based Hybrid Energy System for a Small Community. *Energy* **2021**, *214*, 118977. [[CrossRef](#)]
4. Bahlawan, H.; Morini, M.; Pinelli, M.; Poganietz, W.-R.; Spina, P.R.; Venturini, M. Optimization of a Hybrid Energy Plant by Integrating the Cumulative Energy Demand. *Appl. Energy* **2019**, *253*, 113484. [[CrossRef](#)]
5. Suganthi, L.; Samuel, A.A. Energy Models for Demand Forecasting—A Review. *Renew. Sustain. Energy Rev.* **2012**, *16*, 1223–1240. [[CrossRef](#)]
6. Su, B. Utilizing Hybrid Ensemble Prediction Model in Order to Predict Energy Demand in Sweden: A Machine-Learning Approach. Master's Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2022.

7. Naghibi, S.A.; Ahmadi, K.; Daneshi, A. Application of Support Vector Machine, Random Forest, and Genetic Algorithm Optimized Random Forest Models in Groundwater Potential Mapping. *Water Resour. Manag.* **2017**, *31*, 2761–2775. [\[CrossRef\]](#)
8. Yagli, G.M.; Yang, D.; Srinivasan, D. Automatic Hourly Solar Forecasting Using Machine Learning Models. *Renew. Sustain. Energy Rev.* **2019**, *105*, 487–498. [\[CrossRef\]](#)
9. Shahsavari, A.; Bagherzadeh, S.A.; Afrand, M. Application of Artificial Intelligence Techniques in Prediction of Energetic Performance of a Hybrid System Consisting of an Earth-Air Heat Exchanger and a Building-Integrated Photovoltaic/Thermal System. *J. Sol. Energy Eng.* **2021**, *143*, 051002. [\[CrossRef\]](#)
10. Pino-Mejías, R.; Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A. Comparison of Linear Regression and Artificial Neural Networks Models to Predict Heating and Cooling Energy Demand, Energy Consumption and CO₂ Emissions. *Energy* **2017**, *118*, 24–36. [\[CrossRef\]](#)
11. Yalcintas, M.; Akkurt, S. Artificial Neural Networks Applications in Building Energy Predictions and a Case Study for Tropical Climates. *Int. J. Energy Res.* **2005**, *29*, 891–901. [\[CrossRef\]](#)
12. Wong, S.L.; Wan, K.K.W.; Lam, T.N.T. Artificial Neural Networks for Energy Analysis of Office Buildings with Daylighting. *Appl. Energy* **2010**, *87*, 551–557. [\[CrossRef\]](#)
13. Shiau, Y.-H.; Yang, S.-F.; Adha, R.; Muzayyanah, S. Modeling Industrial Energy Demand in Relation to Subsector Manufacturing Output and Climate Change: Artificial Neural Network Insights. *Sustainability* **2022**, *14*, 2896. [\[CrossRef\]](#)
14. Ahmad, M.W.; Reynolds, J.; Rezgui, Y. Predictive Modelling for Solar Thermal Energy Systems: A Comparison of Support Vector Regression, Random Forest, Extra Trees and Regression Trees. *J. Clean. Prod.* **2018**, *203*, 810–821. [\[CrossRef\]](#)
15. Wang, Z.; Wang, Y.; Zeng, R.; Srinivasan, R.S.; Ahrentzen, S. Random Forest Based Hourly Building Energy Prediction. *Energy Build.* **2018**, *171*, 11–25. [\[CrossRef\]](#)
16. Wei, S.; Yi, C.; Fang, W.; Hendrey, G. A Global Study of GPP Focusing on Light-use Efficiency in a Random Forest Regression Model. *Ecosphere* **2017**, *8*, e01724. [\[CrossRef\]](#)
17. Salem, K.M.; Rey-Hernández, J.M.; Elgharib, A.O.; Rey-Martínez, F.J. Optimizing Energy Forecasting Using ANN and RF Models for HVAC and Heating Predictions. *Appl. Sci.* **2025**, *15*, 6806. [\[CrossRef\]](#)
18. Salem, K.M.; Rady, M.; Aly, H.; Elshimy, H. Design and Implementation of a Six-Degrees-of-Freedom Underwater Remotely Operated Vehicle. *Appl. Sci.* **2023**, *13*, 6870. [\[CrossRef\]](#)
19. Salem, K.M.; Elreafay, A.M.; Abumandour, R.M.; Dawood, A.S. Modeling Two-Phase Gas-Solid Flow in Axisymmetric Diffusers Using Cut Cell Technique: An Eulerian-Eulerian Approach. *Bound. Value Probl.* **2024**, *2024*, 150. [\[CrossRef\]](#)
20. Abumandour, R.M.; El-Reafay, A.M.; Salem, K.M.; Dawood, A.S. Numerical Investigation by Cut-Cell Approach for Turbulent Flow through an Expanded Wall Channel. *Axioms* **2023**, *12*, 442. [\[CrossRef\]](#)
21. Elreafay, A.M.; Salem, K.M.; Abumandour, R.M.; Dawood, A.S.; Al Nuaimi, S. Effect of Particle Diameter and Void Fraction on Gas–Solid Two-Phase Flow: A Numerical Investigation Using the Eulerian–Eulerian Approach. *Comput. Part. Mech.* **2024**, *12*, 289–311. [\[CrossRef\]](#)
22. Salem, K.M.; Rey-Hernández, J.M.; Rey-Martínez, F.J.; Elgharib, A.O. Assessing the Accuracy of AI Approaches for CO₂ Emission Predictions in Buildings. *J. Clean. Prod.* **2025**, *513*, 145692. [\[CrossRef\]](#)
23. Feng, Y.; Yan, Y.; Shi, K.; Zhang, Z. Reducing Carbon Emission at the Corporate Level: Does Artificial Intelligence Matter? *Environ. Impact Assess. Rev.* **2025**, *114*, 107911. [\[CrossRef\]](#)
24. Xie, B.-C.; Wanke, P.F.; Wang, H.; Chen, X.-P. Environmental Assessment of Energy Structure Transition. *Environ. Impact Assess. Rev.* **2026**, *117*, 108181. [\[CrossRef\]](#)
25. Li, J.; Huang, L.; Zhai, J.; Wang, S. Reforming Trade Governance for Sustainable Resource Flows: Ecologically Unequal Exchange in Pan-Eurasia. *Environ. Impact Assess. Rev.* **2026**, *118*, 108247. [\[CrossRef\]](#)
26. Fadoul, F.F.; Hassan, A.A.; Çağlar, R. Integrating Autoencoder and Decision Tree Models for Enhanced Energy Consumption Forecasting in Microgrids: A Meteorological Data-Driven Approach in Djibouti. *Results Eng.* **2024**, *24*, 103033. [\[CrossRef\]](#)
27. Ibeuchi, C.C. Exploring Non-Linear Modes of the Subtropical Indian Ocean Dipole Using Autoencoder Neural Networks. *Environ. Res. Clim.* **2023**, *3*, 011001. [\[CrossRef\]](#)
28. Rey-Hernández, A.; San José-Alonso, J.; Picallo-Perez, A.; Rey-Martínez, F.J.; Elgharib, A.O.; Rey-Hernández, J.M.; Salem, K.M. A Predictive Approach for Energy Efficiency and Emission Reduction in University Campuses. *Appl. Sci.* **2025**, *15*, 9419. [\[CrossRef\]](#)
29. Salem, K.M.; Mohamed, M.S.; ElMessmary, M.H.; Ehsan, A.; Elgharib, A.O.; ElShimy, H. Design and Development of Cost-Effective Humanoid Robots for Enhanced Human–Robot Interaction. *Automation* **2025**, *6*, 41. [\[CrossRef\]](#)
30. Román-Portabales, A.; López-Nores, M.; Pazos-Arias, J.J. Systematic Review of Electricity Demand Forecast Using ANN-Based Machine Learning Algorithms. *Sensors* **2021**, *21*, 4544. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Acikgoz, O.; Çebi, A.; Dalkilic, A.S.; Koca, A.; Çetin, G.; Gemici, Z.; Wongwises, S. A Novel ANN-Based Approach to Estimate Heat Transfer Coefficients in Radiant Wall Heating Systems. *Energy Build.* **2017**, *144*, 401–415. [\[CrossRef\]](#)
32. Ahmed, N.; Ding, Y.; Canti, G.; Anglani, N. Sustainable building's energy management with artificial intelligence and machine learning: A decadal scoping review (2014–2024). *Energy Reports* **2026**, *15*, 109082. [\[CrossRef\]](#)

33. Salem, K.M.; Elgharib, A.O.; Rey-Hernández, J.M.; Rey-Martínez, F.J. A Generalizable Hybrid AI-LSTM Model for Energy Consumption and Decarbonization Forecasting. *Sustainability* **2025**, *17*, 10882. [\[CrossRef\]](#)
34. Farnaaz, N.; Jabbar, M.A. Random Forest Modeling for Network Intrusion Detection System. *Procedia Comput. Sci.* **2016**, *89*, 213–217. [\[CrossRef\]](#)
35. Cody, C.; Ford, V.; Siraj, A. Decision Tree Learning for Fraud Detection in Consumer Energy Consumption. In *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*; IEEE: Piscataway, NJ, USA, 2015; pp. 1175–1179.
36. Ahmad, M.W.; Mourshed, M.; Rezgui, Y. Trees vs Neurons: Comparison between Random Forest and ANN for High-Resolution Prediction of Building Energy Consumption. *Energy Build.* **2017**, *147*, 77–89. [\[CrossRef\]](#)
37. Li, Y.; Wang, R.; Li, Y.; Zhang, M.; Long, C. Wind Power Forecasting Considering Data Privacy Protection: A Federated Deep Reinforcement Learning Approach. *Appl. Energy* **2023**, *329*, 120291. [\[CrossRef\]](#)
38. Zhang, Z.; Li, Y.; Li, Y.; Liu, Y.; Xie, X.; Zeng, Z. A Multi-Task End-to-End Multivariate Long-Sequence Time Series Prediction Model for Load Forecasting. *IEEE Trans. Smart Grid* **2026**, *17*, 715–732. [\[CrossRef\]](#)
39. Cabello, O. *Evaluación de La Demanda Eléctrica Del Sistema de Aire Acondicionado Del Bloque Habitacional Del Hotel Jagua by Meliá Mediante Simulación Energética En Design Building*; Cienfuegos University: Cienfuegos, Cuba, 2023.
40. Balbis Morejón, M.; Cabello Justafré, O.; Cabello Eras, J.J.; Rey-Hernández, J.M.; Rey-Martínez, F.J. Energy Optimization Gaps in Hotel Retrofits for Subtropical Climates. *Sustainability* **2025**, *17*, 5167. [\[CrossRef\]](#)
41. Li, Z.; Dai, J.; Chen, H.; Lin, B. An ANN-Based Fast Building Energy Consumption Prediction Method for Complex Architectural Form at the Early Design Stage. In *Building Simulation*; Springer: Berlin/Heidelberg, Germany, 2019; Volume 12, pp. 665–681.
42. Hu, Q.; So, A.T.P.; Tse, W.L.; Ren, Q. Development of ANN-Based Models to Predict the Static Response and Dynamic Response of a Heat Exchanger in a Real MVAC System. *J. Phys. Conf. Ser.* **2005**, *23*, 110. [\[CrossRef\]](#)
43. Mshragi, M.; Petri, I. FastML-GA: FPGA-accelerated machine learning for real-time energy HVAC optimization in buildings. *Neural Comput. Appl.* **2026**, *38*, 25. [\[CrossRef\]](#)
44. Fawagreh, K.; Gaber, M.M.; Elyan, E. Random Forests: From Early Developments to Recent Advancements. *Syst. Sci. Control. Eng. Open Access J.* **2014**, *2*, 602–609. [\[CrossRef\]](#)
45. Chen, Y.-T.; Piedad, E., Jr.; Kuo, C.-C. Energy Consumption Load Forecasting Using a Level-Based Random Forest Classifier. *Symmetry* **2019**, *11*, 956. [\[CrossRef\]](#)
46. Demir, E.; Gunal, S. Short-term electricity consumption forecasting with deep learning: E. Demir, S. Gunal. *J. Supercomput.* **2025**, *81*, 1108. [\[CrossRef\]](#)
47. Abbasimehr, H.; Paki, R.; Bahrini, A. A Novel XGBoost-Based Featurization Approach to Forecast Renewable Energy Consumption with Deep Learning Models. *Sustain. Comput. Inform. Syst.* **2023**, *38*, 100863. [\[CrossRef\]](#)
48. Zhu, Z.; He, K. Prediction of Amazon's Stock Price Based on ARIMA, XGBoost, and LSTM Models. *Proc. Bus. Econ. Stud.* **2022**, *5*, 127–136. [\[CrossRef\]](#)
49. Bitirgen, K.; Filik, Ü.B. Electricity Price Forecasting Based on Xgboost and Arima Algorithms. *BSEU J. Eng. Res. Technol.* **2020**, *1*, 7–13.
50. George, J.; Yadav, J.; Nair, A.M.; Peter, M.V.; Alapatt, B.P.; Baby, R. Improving Groundwater Forecasting Accuracy with a Hybrid ARIMA-XGBoost Approach. In *2024 3rd International Conference for Advancement in Technology (ICONAT)*; IEEE: Piscataway, NJ, USA, 2024; pp. 1–7.
51. de Jesús Rubio, J.; Garcia, D.; Sossa, H.; Garcia, I.; Zacarias, A.; Mujica-Vargas, D. Energy Processes Prediction by a Convolutional Radial Basis Function Network. *Energy* **2023**, *284*, 128470. [\[CrossRef\]](#)
52. Lin, W.-M.; Gow, H.-J.; Tsai, M.-T. An Enhanced Radial Basis Function Network for Short-Term Electricity Price Forecasting. *Appl. Energy* **2010**, *87*, 3226–3234. [\[CrossRef\]](#)
53. Ghods, L.; Kalantar, M. Long-Term Peak Demand Forecasting by Using Radial Basis Function Neural Networks. *Iran. J. Electr. Electron. Eng.* **2010**, *6*, 3.
54. El Mokhtari, K.; McArthur, J.J. Autoencoder-Based Fault Detection Using Building Automation System Data. *Adv. Eng. Inform.* **2024**, *62*, 102810. [\[CrossRef\]](#)
55. Chen, S.; Guo, W. Auto-Encoders in Deep Learning—A Review with New Perspectives. *Mathematics* **2023**, *11*, 1777. [\[CrossRef\]](#)
56. Fan, C.; Xiao, F.; Zhao, Y.; Wang, J. Analytical Investigation of Autoencoder-Based Methods for Unsupervised Anomaly Detection in Building Energy Data. *Appl. Energy* **2018**, *211*, 1123–1135. [\[CrossRef\]](#)
57. Ramos, D.; Faria, P.; Morais, A.; Vale, Z. Using Decision Tree to Select Forecasting Algorithms in Distinct Electricity Consumption Context of an Office Building. *Energy Rep.* **2022**, *8*, 417–422. [\[CrossRef\]](#)

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.