Multi-Timescale Energy Consumption Management in Smart Buildings Using Hybrid Deep Artificial Neural Networks.

IBUDE, Favour, OTEBOLAKU, Abayomi <http://orcid.org/0000-0002-4320-9061>, AMEH, Jude <http://orcid.org/0009-0001-4523-5204> and IKPEHAI, Augustine

Published version

IBUDE, Favour, OTEBOLAKU, Abayomi, AMEH, Jude and IKPEHAI, Augustine (2024). Multi-Timescale Energy Consumption Management in Smart Buildings Using Hybrid Deep Artificial Neural Networks. Journal of Low Power Electronics and Applications, 14 (4): 54. [Article]

Copyright and re-use policy

Journal of
Low Power Electronics and Applications

MDPI

Article

Multi-Timescale Energy Consumption Management in Smart Buildings Using Hybrid Deep Artificial Neural Networks

Favour Ibude [1], Abayomi Otebolaku [1,*], Jude E. Ameh [1] and Augustine Ikpehai [2]

1 School of Computing and Digital Technologies, Sheffield Hallam University, Sheffield S1 2NU, UK; favour.ibude@outlook.com (F.I.); j.e.ameh@shu.ac.uk (J.E.A.)
2 School of Engineering and Built Environment, Sheffield Hallam University, Sheffield S1 2LX, UK; a.ikpehai@shu.ac.uk
* Correspondence: a.otebolaku@shu.ac.uk

Abstract: Demand side management is a critical issue in the energy sector. Recent events such as the global energy crisis, costs, the necessity to reduce greenhouse emissions, and extreme weather conditions have increased the need for energy efficiency. Thus, accurately predicting energy consumption is one of the key steps in addressing inefficiency in energy consumption and its optimization. In this regard, accurate predictions on a daily, hourly, and minute-by-minute basis would not only minimize wastage but would also help to save costs. In this article, we propose intelligent models using ensembles of convolutional neural network (CNN), long-short-term memory (LSTM), bi-directional LSTM and gated recurrent units (GRUs) neural network models for daily, hourly, and minute-by-minute predictions of energy consumptions in smart buildings. The proposed models outperform state-of-the-art deep neural network models for predicting minute-by-minute energy consumption, with a mean square error of 0.109. The evaluated hybrid models also capture more latent trends in the data than traditional single models. The results highlight the potential of using hybrid deep learning models for improved energy efficiency management in smart buildings.

Keywords: smart buildings; energy consumption; hybrid deep learning; energy forecasting; building energy management systems

Citation: Ibude, F.; Otebolaku, A.; Ameh, J.E.; Ikpehai, A. Multi-Timescale Energy Consumption Management in Smart Buildings Using Hybrid Deep Artificial Neural Networks. *J. Low Power Electron. Appl.* **2024**, *14*, 54. https://doi.org/10.3390/jlpea14040054

Academic Editor: Chih-Hung Chen

Received: 30 August 2024
Revised: 4 October 2024
Accepted: 30 October 2024
Published: 7 November 2024

1. Introduction

Achieving efficiency is an important focus in the energy sector. The global energy crisis has further emphasized the consideration for efficient energy consumption [1]. In addition, the new wave of energy crises in the UK and Europe, triggered by Russia's invasion of Ukraine in the last couple of years, as well as the increase in world population, has led to the rapid increase in energy demand and higher costs. Therefore, the need to reduce greenhouse emissions, extreme weather conditions, etc., all of which have impacted the surge in demand for energy [2], has prompted many countries to embark on campaigns to minimize energy wastage [3]. According to the World Energy Council, climate change has been one of the biggest challenges affecting all regions of the world. For instance, European Climate Action listed the negative impacts of this and aims to minimize greenhouse emissions and improve energy efficiency by reducing energy consumption. The biggest energy consumers according to [4] are buildings, contributing most of the total energy usage and carbon emissions in the world. Similarly, energy demands are projected to increase by 55% from 2005 to 2030, with buildings accounting for 40% of the total energy consumed [5]. Due to this significant challenge, more attention is now paid to smart buildings by providing comfortable, economical, and sustainable operations for occupants.

In recent years, emerging and disruptive technologies such as artificial intelligence (AI) and the Internet of Things (IoT) have been shaping the energy management future, building a world of smart and connected agents that require minimal or no human intervention. These technologies are being integrated into building automation systems to develop smarter buildings [6]. Smart buildings have been widely adopted by developed countries due to the popularity of intelligent technologies such as smart grids [7] and their ability to support sustainable and efficient energy management systems (EEMSs). EEMSs are the main feature of smart buildings for managing energy use, hence the need (as an integral part of the EEMS) for accurate energy consumption predictions to aid occupants in managing, planning, and minimizing energy waste and cost [8].

Due to the current global increase in energy prices, smart buildings have become increasingly popular because of their inbuilt sensors which help to monitor occupants' behaviours and optimize their energy consumption. Currently, the UK government policy advises installing sustainable technologies in houses. Therefore, prediction of smart building energy use is an important factor for efficient management of energy consumption. However, the main challenge in smart building energy management systems is poor prediction performance [9], particularly false positive or false negative predictions.

Therefore, accurate energy consumption prediction in smart buildings is essential and it represents an important challenge for efficient set points of critical loads, such as heating, ventilation and air conditioning (HVAC) and scheduling of energy-production assets. Studies have demonstrated that predicting the consumption of each appliance will improve occupants' attitudes towards energy saving [10]. The ability to forecast energy consumption at periodic intervals can help building occupants to anticipate and adjust the operation of their appliances and equipment, thus leading to improved energy efficiency management, sustainable development, reductions in energy costs, improved environmental influence, and reductions in energy expenses [11]. Energy from power plants is instantly consumed as it is generated; hence, accurate forecasting of energy consumption will support the stability and continuous improvement of power supply.

Furthermore, home energy management has received considerable attention in recent times because of the need for energy consumers to minimize the overall electricity consumption as the cost of living continues to increase. Therefore, more technology-based approaches are being explored by researchers to automate energy management. One technology-based approach is the use of artificial intelligent models. Several intelligent approaches, such as mathematical models and classical machine learning models, for energy consumption prediction, have been explored in the past by researchers. However, approaches based on deep neural networks are generally considered to produce better forecasting performance and robustness than classical methods [12–14], minimizing both false positive and false negative predictions.

Thus, several researchers have applied deep neural networks for forecasting and predicting energy consumptions [1,6–8]. For example, recently, Mubarak et al. [15] explored hybrid deep learning models consisting of long short-term memory and self-attention (LSTM–Attention), incorporating explicit time encoding to forecast one-hour-ahead active and reactive power usage. In addition, Sunder et al. [16] also proposed hybrid models consisting of a combination of various deep learning models such as transformer-based models, graph neural networks, recurrent neural networks, e.g., BiLSTM with improved CNN models for accurate long-term predictions of energy load in smart buildings. These research works and many more have demonstrated that a combination of deep neural network models could produce better prediction performances than individual models. Thus, this article presents the use of hybrid models that integrate convolutional neural networks (CNNs), gated recurrent units (GRUs), long short-term memory (LSTM), and CNN–Bidirectional LSTM to predict daily, hourly, and minute-by-minute consumptions of energy, unlike other works that are focused on long-term predictions of energy consumptions.

The key contributions of this article are threefold:

- Providing a comprehensive review of existing energy prediction approaches in smart buildings.
- Investigating the use of hybrid convolutional neural networks, CNN–GRU, CNN–LSTM, and CNN–Bidirectional LSTM, for predicting smart buildings' energy consumption on minute-by-minute, hourly, and daily basis, relying on the inherent strengths of

these individual models for recognizing temporal and spatial relationships in the data for accurate predictions of energy consumption.

- Extensive performance comparisons of the models with other state-of-the-art deep learning models for predicting energy consumption at the mentioned time intervals.

The remainder of this article is structured as follows:

In Section 2, a review of related works focusing on aspects covered by the proposed solution is presented. Section 3 presents the methodology of the proposed solution. In Section 4, experimental validation of the proposed solution and results are presented. Section 5 analyses and discusses the results and their significance. Finally, in Section 6, we conclude the article and present recommendations for future work.

2. Related Works

Predicting energy usage is an important factor for achieving sustainable energy efficiency [17]. Over the last decade, increasing energy demand has inspired researchers to find the best approaches for minimizing energy consumption, reducing the cost of energy, making informed energy decisions, and improving energy utilization.

To achieve these aims, several methods have been explored for energy consumption predictions. These prediction models can be generally placed into four categories: statistical models, classical machine learning models, deep learning models, and hybrid models.

These statistical methods use historical data to develop statistical models to analyse and estimate future energy consumption [18–20]. Regression models are the most common statistical methods used for predicting energy consumption. Linear, multivariate, and other regression models have been widely used [18,19]. Other examples include the autoregressive integrated moving average model (ARIMA) [20] and the support vector regressor (SVR) approach [14,21]. The authors of [22] utilized multiple regression models to predict energy demands of heating systems in residential buildings. The model incorporates various features that influence energy consumption including the global heat loss coefficient, south equivalent surface, indoor set-point and sol-air temperature differences. A good accuracy of 0.987 correlation coefficient was achieved highlighting the effectiveness of the model for forecasting heating energy demands, based on the identified features. In [23], the researchers applied linear regression to estimate energy consumption in an institutional building. This study demonstrates the efficacy of a regression model for energy forecasting, but it only performed well when dealing with smaller datasets.

Another category comprises researchers who explored classical machine learning methods. Several traditional machine algorithms have been explored by researchers to predict energy consumption [7,9,12,14,24]. Machine learning allows systems to learn automatically and improve by observing patterns in the data. In this regard, several classical machine learning algorithms and techniques are being adopted for predicting energy consumption. In [25], for example, a study was conducted to analyse and predict the energy consumed in smart buildings in Malaysia. In that study, hourly consumption data were collected from commercial smart buildings. K-nearest neighbour (KNN), a popular classical machine learning algorithm, was trained for energy consumption prediction. The result shows high accuracy with k = 5. Although KNN performs well with a small quantity of data, it requires huge computational resources when handling large datasets because it memorises the entire data. In [26], the researchers investigated the use of ANN and SVM and compared their performances with that of KNN. The results demonstrate that, even though ANN and SVM are more complex models compared to KNN, SVM demonstrates a better performance than KNN when analysing and predicting energy consumption.

Another study [27] introduced a forecasting method that implements a two-stage hybrid approach for short-term load forecasting. The first stage explored time-series methods whilst in the second stage they enhanced the performance of time-series methods by analyzing their deviations using techniques such as linear regression, quadratic programming, and support vector machines (SVM). However, SVM appeared to be unsuitable for large datasets because the bigger the dataset, the longer the linear training time. In addition, the results from previous machine learning studies show that every method performed better or worse depending on factors such as the size of the dataset used, the data preprocessing approach, and the duration of training [28]. Another example is the work in [11], where a multilayer perceptron algorithm was trained to forecast heating and cooling loads in a residential building, achieving a good performance. The researchers in [29] used a deep residual neural network to forecast electrical energy consumption, providing day-ahead estimations in a residential building; the forecast was tested individually on several residential buildings, and the model obtained some good results with an error rate of 8% for hourly forecasting(for 22 kwh) and 2% error for daily forecasting(of 131 kwh), which were better than that achieved by the benchmark model.

Researchers and developers alike are exploiting deep neural networks for solving more complex problems that the classical machine learning algorithms struggled to solve [30]. Thus, apart from the use of classical machine learning algorithms for energy consumption prediction, such as KNN, SVM, ANN, and random forest, there have been recent developments and successes in the use of deep learning models for solving complex forecasting problems based on time–series data [1,6–8,15,16,28,31]. In addition, CNNs have achieved ground-breaking results in computer vision problems, with their capability to extract latent spatial features from data. In this regard, several deep learning models such as CNN, LSTM, and their variants—transformer models, etc.—are being widely explored for energy consumption forecasting. For example, the authors of [32] used various deep learning methods such as GRU, LSTM, and RNN to predict the energy consumption of smart buildings. The results demonstrate better performance than those obtained from classical machine learning algorithms, with improvements in prediction accuracy.

LSTM, Bidirectional LSTM, and GRU, for example, may exhibit better performances in extracting temporal sequences in multiple time steps from time–series data. However, CNNs do not have this capability; rather, they have the capability to extract spatial features and local patterns, which LSTM-based models such as the bidirectional LSTM or GRU do not have.

Therefore, attention has recently been shifting to combining multiple deep neural network models as hybrid architectures for solving complex problems. This has the synergistic benefits of hierarchical feature extractions and improved performances, particularly in applications that involve time–series forecasting such as energy consumption predictions. Additionally, existing works have proven that combining these architectures allows for a more comprehensive understanding of hidden patterns in time–series data, leveraging the strengths of the individual model to achieve a better forecasting performance. For example, a novel multi-channel and multi-scale convolutional neural network–long short-term memory (MCSCNN-LSTM) hybrid was presented in [24] to predict energy consumption. The authors confirmed that the hybrid model has a better performance when predicting irregular trends and patterns of energy consumption than when using each architecture alone. There are several other recent works that explore hybrids of various deep learning models to achieve improved performances [13,33–36].

However, in this article, we investigate the efficacy of some of these hybrid models to accurately capture temporal trends in energy consumption data. We explore and evaluate the capacity of combinations of CNN, LSTM, bidirectional LSTM, and GRU to capture sequence patterns at different time intervals, such as minute-by-minute, hourly, and daily intervals, for forecasting energy consumptions.

3. Methodology

In this section, we present the approach and methods explored using the proposed solution for predicting energy consumption in smart buildings.

3.1. Proposed Hybrid Framework for Energy Consumption Predictions in Smart Buildings

Figure 1 illustrates, at a high level, the proposed hybrid method for energy consumption prediction. The figure shows different aspects of the proposed solution, such as data collection, data preprocessing, model training, and evaluation. The spatial and local features are automatically extracted by the CNN convolution layer. The LSTM, bidirectional LSTM, and GRU are explored for capturing temporal patterns in multiple time steps of energy consumption datasets. The models are analysed and evaluated using relevant performance metrics, which are presented in Section 4.3.

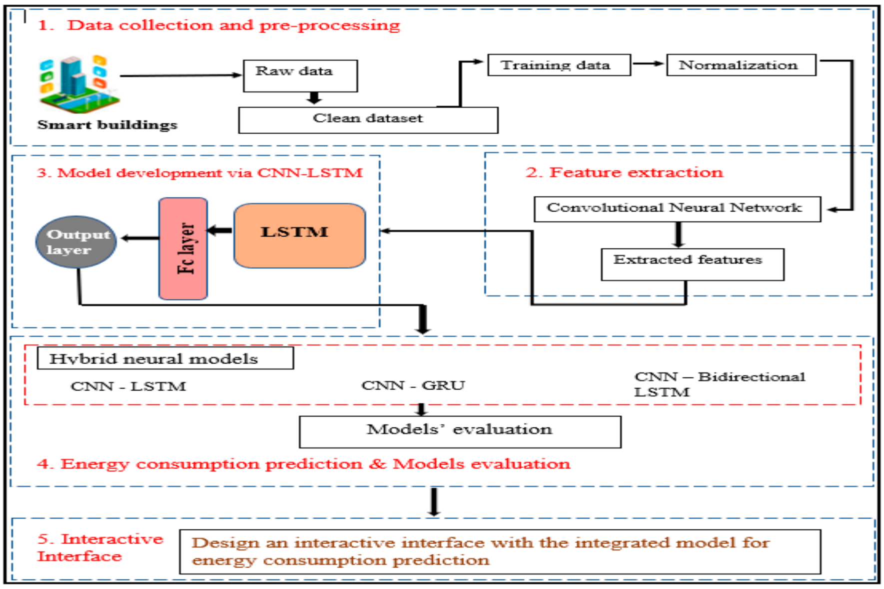

Figure 1. The proposed framework for energy consumption prediction using hybrid deep neural networks.

3.2. Theoretical Overview of the Proposed Solution

In this section, the theoretical underpinnings of our methods, as illustrated in Figures 1–3, are introduced.

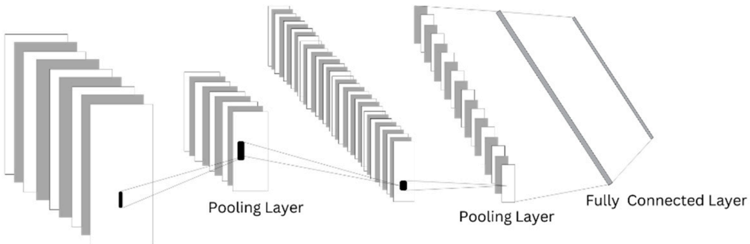

Figure 2. Schematic diagram showing a typical CNN model with LeNet architecture [37].

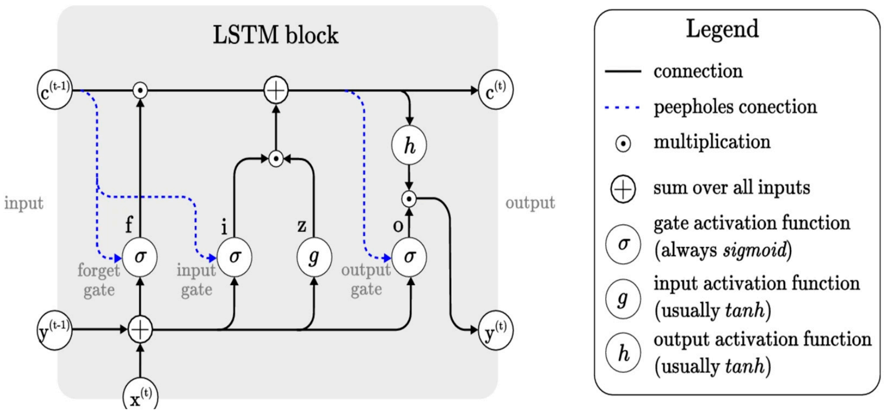

Figure 3. Schematic diagram of LSTM memory block [38].

3.2.1. Convolutional Neural Network Building Blocks (Layers)

A convolutional neural network, as illustrated in Figure 2, consists of multiple layers of architecture known as building blocks. CNNs can transform and process time–series data using three layers, namely the convolutional layer, the pooling layer, and the fully connected layer [30,39].

a. The Convolutional Layer

In CNNs, two main operations are performed, namely linear convolution and non-linear transformations [37,39]. The convolution is a specialized linear operation that uses several kernels. The purpose of this operation is to automatically extract discriminatory spatial features from the raw data. This process involves an element-wise product of the kernel and the input matrix to generate a feature map. After the convolution operation, a non-linear activation function, such as an ReLU (rectified linear unit), is used to transform the output, as shown in Equation (1).

$$ReLU = f(x) = max(0, x) \tag{1}$$

This non-linearity helps the network to capture more complex spatial patterns and relationships, making the CNN better equipped for automatic feature extraction [39]. Assuming that $x_i^a = [x_1, \ldots, x_N]$ are inputs from the power consumption data of each unit, a represents the time intervals (hourly, minute-by-minute, or daily). The output, which depends on the number of convolutional layers in the architecture, of the lth convolutional layer is computed using Equation (2).

$$z_i^{l,j} = \sigma(\textstyle\sum_{k=1}^{K} w_k^j x_{i+k-1}^{l-1,j} + b_j^l) \tag{2}$$

where $w_k^{l,j}$ and b_j^l are weight and bias of the j-th term of the l-th layer; $x_{i+k-1}^{l,j}$ is the input patch; l is the index of the current layer, and σ is the activation function; k represents the size of the filter/kernel.

The activation function $a_j^l = \sigma(z_i^{l,j})$ introduces non-linearity to the CNN layer for detecting the non-linear features of the raw sensing data.

b. The pooling Layer

The pooling layer is important in a convolutional neural network, and it is also called the down-sampling layer; it helps to control overfitting, and in maintaining a translational invariant feature in the model. Pooling layers aid in reducing the size of the model layer feature map, the number of parameters, the computational requirements of the network, and the memory footprints needed to speed up the training process while decreasing the neurons in networks and extracting important features. In the pooling layer, there are different pooling techniques like max polling, min pooling, gated pooling, average pooling, tree pooling, etc. Average pooling and max pooling are the most-used techniques in the pooling layer. Max pooling selects the maximum output value (as shown in Equation (3)) from all units, while average pooling computes the average output from all the windows. Max pooling down-samples the weights assigned to the kernel in the convolutional layer, therefore reducing the possibility of overfitting as well as the computational cost [30,37–39]. In this study, we used the max pooling method because of its robust performance with sparse features and its capability to eliminate unimportant features.

$$f_i^{l,j} = max_{s \in S}(z_{i*T+s}^{l,\,j}) \tag{3}$$

where S is the pooling size, and its stride is denoted by T. $z_{i*T+s}^{l,\,j}$ is the value of the $i - th\ node\ in\ layer\ l$.

c. Fully Connected (FC) Layer

After the pooling process is complete, the resulting feature map is flattened into a one-dimensional vector of features $f^l = [f_1, \ldots, f_l]$, where l is the number of nodes in the last pooling layer. The flattened features are then passed to the fully connected layers, which are also called the dense layers. In this layer, there is a full connectivity of neurons in the first and last layers. This layer helps outline the representation between the input and output.

3.2.2. Long Short-Term Memory Approach

LSTM is a type of recurrent neural network that learns the hidden relationships and patterns between data points in sequence and has contributed widely to deep learning success stories [40–43]. It was developed to handle long-term memory tasks, like speech recognition [34], music generation [41], and energy consumption prediction and forecasting [32,34]. Also, LSTM models can be trained with historical time–series data to make predictions for the future energy consumption of buildings.

About a decade ago, long short-term memory models gained popularity in the building energy consumption prediction and forecasting domain; they have since been used more frequently with other deep learning models such as CNN, and DNN. The tremendous success of LSTM models is due to their capability to solve time–series tasks and memorize information for a longer time steps in networks [43]. They also have the capability to reduce the exploding and vanishing gradient problems that are associated with traditional recurrent neural networks [43,44].

3.2.3. Long Short-Term Memory Block

The LSTM building blocks comprise memory cells that are self-connected, as illustrated in Figure 3. The LSTM cells have the capacity to remember their past states. This is possible using three gates, namely the input gate, the output gate, and the forget gate, to store information over a long-time step [43]. The input gate decides what information will be included and updated in the current timestamp for future prediction. The forget gate decides on the extent to which the information should be remembered or forgotten from previous time steps, whilst the output gate determines the future predicted values. The operation of the three gates can be mathematically expressed as shown in Equations (4)–(9) [37].

$$f^{(t)} = \sigma(W_f \left[h^{(t-1)}, x^t\right] + b_f) \tag{4}$$

$$i^{(t)} = \sigma(W_i \left[h^{(t-1)}, x^t\right] + b_i) \tag{5}$$

$$o^{(t)} = \sigma(W_o \left[h^{(t-1)}, x^t\right] + b_o) \tag{6}$$

$$\tilde{c}^{(t)} = \tanh(W_c \left[h^{(t-1)}, x^t\right] + b_c) \tag{7}$$

$$c^{(t)} = f^{(t)} * c^{(t-1)} + i^{(t)} * \tilde{c}^{(t)} \tag{8}$$

$$h^{(t)} = o^{(t)} * \tanh(c^{(t)}) \tag{9}$$

In LSTM cells, the forget gate (Equation (4)) determines the amount of information from the previous cell state that should be retained or discarded. It takes the hidden state from the previous time step and the current input as inputs. These inputs are passed through a sigmoid function to produce a value between 0 and 1, which modulates the contribution of the previous cell state to the current one.

The input gate (Equation (5)) decides how much new information from the current input should be added to the cell state. It produces a gate value that regulates the contribution of the candidate cell state to the overall cell state.

The output gate (Equation (6)) governs the generation of the hidden state at the current time step, which serves as the output of the LSTM cell. It produces a value that modulates the impact of the updated cell state on the hidden state.

The candidate cell state (Equation (7)) represents the new information that could potentially be added to the current cell state. It is computed by applying the hyperbolic activation function to the weighted sum to produce values between −1 and 1. The input gate controls how much of this candidate state is used in updating the overall cell state.

The updated cell state (Equation (8)) is a combination of the previous cell state, scaled by the forget gate, and the candidate cell state, modulated by the input gate. This equation allows the LSTM to maintain long-term dependencies by selectively incorporating old and new information, depending on the context at each time step.

In summary, Equations (4)–(6) describe the forget, input, and output gates. Equations (7) and (8) use the input and the output gates of the candidate cells to compute the values for the new cell. Equation (9) ensures that the output values of $h^{(t)}$ are always in the interval (−1,1). On the one hand, if the value of the output gate is close to 1, then the gate allows the memory cell of the internal state to impact the subsequent layers. On the other hand, if the value is close to 0, then it prevents the current memory cell from impacting other layers of the network in the current time step.

3.3. The Hybrid CNN-LSTM Method

Figure 4 shows the high-level architecture of the CNN-LSTM method for energy consumption prediction in smart buildings. This study uses a real energy consumption dataset from a smart building. Spatial factors associated with the time–series variables which are multivariate are extracted from the CNN convolution layer and the pooling layers and fed into LSTM layers with outliers removed. The LSTM layer uses transmitted spatial characteristics to model irregularly in the time–series data, such as irregular time patterns, trends, and seasonality. Lastly, the CNN-LSTM model can produce predicted energy consumption in a fully connected state. The predicted values of energy consumption are then analysed and evaluated using several evaluation metrics, which are presented in Section 4.3.

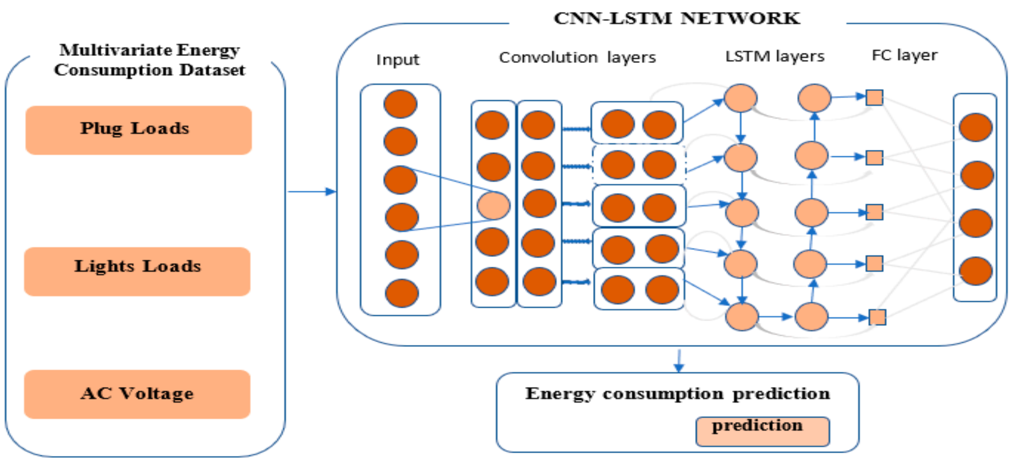

Figure 4. The high-level architecture of the models showing one of the hybrid models, i.e., the CNN-LSTM model.

Predicting energy consumption using the CNN-LSTM methods requires a series of connections between the CNN and the LSTM. The model can keep irregular time information and extract complicated hidden features from the building's sensor data collected to predict energy consumption trends. First, the upper layers of the CNN-LSTM model architecture consist of one or more CNN layers. This CNN layer can receive (from the input layer) various features that define the energy consumption of appliances, building characteristics, the occupants' behaviours, household occupancy, weather information, seasons of the year, and time. The CNN layer is responsible for automatic feature extraction from the data it receives from the input layer. The extracted features from the CNN layer are then fed to the LSTM layer. Meanwhile, the hidden layer of the architecture is the heart of the network, where the information processing mechanisms such as feature extraction, regularization etc. happen. The convolutional layers, the pooling layers, the LSTM layers dropout layers, and ReLU layers, also known as the activation function, all constitute the hidden layers of the architecture. In the convolution layer, the convolution operation is applied to the incoming sequence of the time–series features; then, the result is passed into the next layer. Visual stimulation of individual neurons is emulated by the convolution operation. The individual neurons in the convolution layer then process only the multivariate data for the receiving field, thus reducing parameters. The dropout layer is added for regularization to allow the model to generalize, i.e., to prevent overfitting and to improve overall performance of the model.

LSTM is the lower layer of the CNN-LSTM model that memorizes time information regarding significant features from the energy consumption sensor extracted from the CNN. LSTM can remember long-term information by updating the hidden state, which makes it easy to understand the temporal relationship. The obtained output value from the CNN layer is passed into the LSTM gate units. LSTM is best for predicting energy consumption because it solves the issues posed by vanishing and explosive gradients, which are associated with RNNs. LSTMs are made up of memory cells that update their current state using each gate unit activation function. The activation function is a continuous value from 0 to 1 and it is controlled to fit into the value.

$$i^t = \delta(W_{Pi}p_t + W_{hi}h_{t-1} + W_{ci}oc_{t-1} + b_i) \tag{10}$$

$$f^t = \delta(W_{Pf}p_t + W_{hf}h_{t-1} + W_{cf}oc_{t-1} + b_f) \tag{11}$$

$$o^t = \delta(W_{Po}p_t + W_{ho}h_{t-1} + W_{co}oc_t + b_o) \tag{12}$$

where $\mathrm{h_t}$ is the LSTM cell hidden state, which is updated in every step t. Equations (10)–(12) above show the individual gate unit (input, output, and forget gates) operations that constitute the LSTM; notations *i*, *f*, and *o* represent the outputs of the individual gates.

$$c^t = f^t \ o \ c_{t-1} + i^t \ \mathrm{o} \ \delta(W_{Pc}p_t + W_{hc}h_{t-1} + b_c) \tag{13}$$

$$h^t = o^t o \delta(c^t) \tag{14}$$

In the above Equations (13) and (14) [36], the notation *c* and *h* are the hidden states. Also, for the cell states that are determined through the gate units, the activation function—such as Tanh—is represented with the o notation; please also note that the non-linear activation function confines the input into the −1,1 range. The bias vector is the b notation, while w represents each gate unit weight matrix. The notation $\mathrm{p_t}$ stores complex features as output and the output is used in the LSTM memory cells as an input.

In addition to the unidirectional LSTM model, two other variants of LSTM—namely the bidirectional LSTM and gated recurrent units (GRUs)—were also trained and evaluated in this article to compare their performance for minute-by-minute, hourly, and daily energy consumption forecasting.

The BiLSTM consists of two independent LSTM layers for processing the input data in two opposing directions. One layer is used for processing the data in the forward direction, like the normal LSTM, whereas the other layer processes the data in the opposite direction. This works by dividing and connecting the neurons of the normal LSTM for the backward states (the negative time direction) and the other does this for the forward states (the positive time directions) [44].

In the GRU architecture, the three gates of the LSTM are replaced with two gates, namely the reset and the update gates. The reset gate controls how much of the previous state is remembered, whereas the update gate controls how much of the new state is a copy of the old state [44].

The fully connected layer, which is the last layer, consists of densely connected neurons. The fully connected layers receive inputs from the LSTM, BiLSTM, or GRU layers to produce the final output. The inputs are flattened to produce a one-dimensional vector feature before they are fed into the fully connected layers [35].

4. Experiments and Results

In this section, we present an experimental evaluation of the proposed solutions. A total of 12 models, including LSTM, CNN, GRU, and hybrid models (including CNN-LSTM, CNN-GRU, and CNN–Bidirectional LSTM) were trained and evaluated for their ability to accurately predict energy consumption in smart buildings. The experiments were conducted on minute-by-minute, hourly, and daily time intervals, and the model performances were measured using the root mean square error (RMSE), the mean absolute error (MAE), and the mean squared error (MSE) metrics.

The proposed CNN-LSTM model was compared to other models in terms of its ability to make short-term, medium-term, and long-term predictions. The dataset was aggregated from minute-by-minute timestamps to hourly and daily timestamps to evaluate the model performance at different time intervals. A sliding-window algorithm with a window size of 2 was utilized in the experiment, where the model was fed with two consecutive time steps as input and used to predict the next value.

4.1. Energy Consumption Dataset and Analysis

In this study, the energy consumption dataset of a seven-story office building in Thailand was adopted and used for the experiments, as illustrated in Figure 5 [45]. Several time–series variables of the dataset were used to predict the plug load energy consumption. This dataset is displayed as a one-minute time unit with real energy consumption. A total of 790,558 datapoints from 2018 to 2019, specifically collected from only the first floor of the office buildings, were used to train the models. In addition, a total of 49,456 missing data points were identified. Table 1 describes each variable contained in the dataset. Seasonality is usually represented by periodic trends going up and down in the dataset; trends are patterns in the data that span across the seasonal period, while residual is the noise present in the data that cannot be explained. Figure 6 shows the trends in seasonality and residual energy consumption for one of the plug loads, represented as variable “z1_plug”.

	z1_light	z1_plug	z2_AC1	z2_AC2	z2_AC3	z2_AC4	z2_light	z2_plug	z3_light	z3_plug	z4_light
Date											
2018-07-01 00:00:00	12.94	18.56	45.24	0.01	0.01	0.00	13.76	17.64	10.92	0.89	35.76
2018-07-01 00:01:00	12.97	18.55	45.28	0.02	0.01	0.01	13.76	17.21	10.95	0.87	35.81
2018-07-01 00:02:00	12.97	18.55	45.24	0.01	0.01	0.01	13.79	17.18	10.94	0.86	35.78
2018-07-01 00:03:00	12.98	18.58	45.26	0.02	0.01	0.00	13.81	16.64	10.94	0.85	35.83
2018-07-01 00:04:00	13.01	18.60	45.22	0.02	0.01	0.01	13.83	15.69	10.97	0.85	35.86

Figure 5. An overview of the energy consumption dataset. The electricity consumption dataset is of individual air conditioning units, lighting, and plug loads in each of the 33 zones of the 33 zones of the building [45].

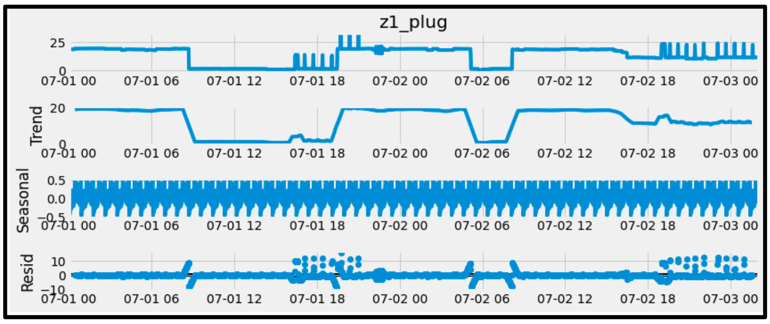

Figure 6. Insights from the time–series dataset showing trends and seasonality of energy consumption for a single plug load.

Table 1. Description of features in the dataset with eleven (11) variables.

Variable Names	Description
Date	The electricity/energy consumption was recorded on the first floor of the office buildings with time-series data for year, month, day, hour, minute, and second
z1_light	Power consumption of lighting load for zone 1 (kW)
z1_plug	Power consumption of plug load for zone 1 (kW)
z2_AC1	Power consumption of AC unit 1 (kW)
z2_AC2	Power consumption of AC unit 2 (kW)
z2_AC3	Power consumption of AC unit 3 (kW)
z2_AC4	Power consumption of AC unit 4 (kW)
z2_light	Power consumption of lighting load for zone 2 (kW)
z2_plug	Power consumption of plug load for zone 2 (kW)
z3_light	Power consumption of lighting load for zone 3 (kW)
z3_plug	Power consumption of plug load for zone 3 (kW)
z4_light	Power consumption of lighting load for zone 4 (kW)

4.1.1. Energy Consumption by Month

As part of the initial analysis of the datasets, we compared the trends of energy consumption during the summer and winter months. From Figure 7a,b, the trends show that, from June to August, the energy consumed was below 40 KW. This could be because, during summer, people hardly turn on the heaters that consume more energy, compared to energy usage in the wintertime, when energy consumption is generally higher than 40 KW, with more plug load.

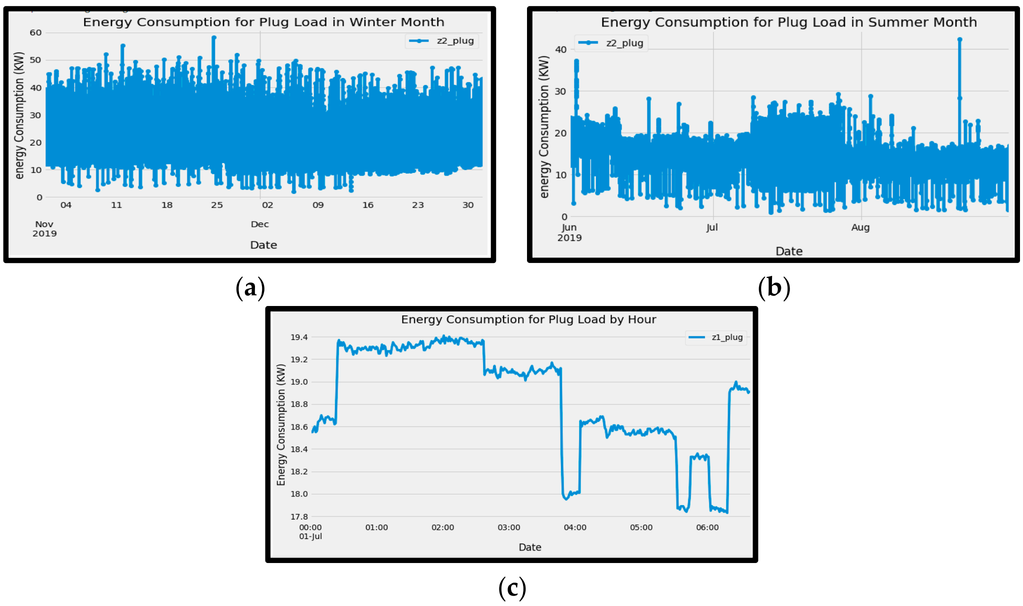

Figure 7. Analysis of experimental dataset. (**a**,**b**) The energy consumption trends from November to December (Winter) and from July to August (Summer). (**c**) The energy consumption trends at different hours of the day.

4.1.2. Energy Consumption by Hour

Figure 7c illustrates the energy consumed per hour; the period from 12:00 pm to 2:30 pm recorded the most energy consumed. From 3:30 pm to 4 pm, there was a drastic drop in consumed energy; this could be because the afternoon work shift was over, and staff had to turn off all loads. From 4 pm to 5:30 pm, there was a peak in energy consumption, this could mean that evening staff had resumed their work in the office.

4.2. Model Development and Training

In this section, we present details for the development and training of the hybrid models.

4.2.1. Data Segmentation

Before feeding the raw energy dataset into the models for training, the data were segmented using a temporal sliding-window algorithm [46]. We applied the sliding-window algorithms to address the time-sensitive nature of the dataset and to aid the computational efficiency (reducing processing time and energy consumption) of the models. In this process, we used a sliding window size of two (window size = 2). This means that the model takes a 2-time step (mins, hours, days, or months) into the future to make a prediction of energy consumption for the third minute, hour, day, or month.

4.2.2. Split the Data into Training, Validation, and Testing

Once the data are segmented through the sliding window process, the data are split into training, validation, and test sets. As can be seen in Figure 8, 70% of the data were used for training, 10% were used for validation, and 20% for testing.

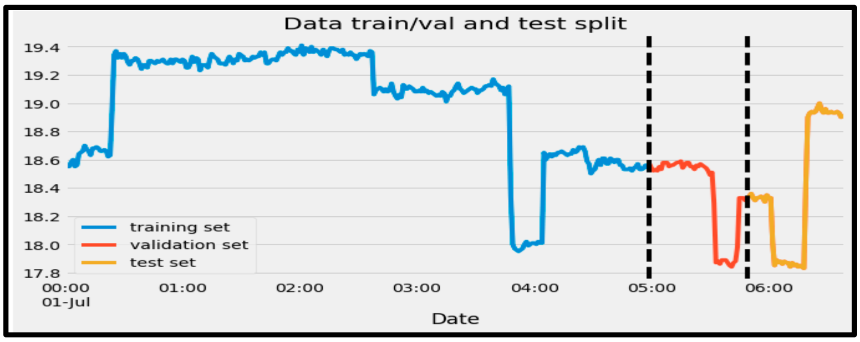

Figure 8. Splitting of data into training and test sets, the data were split by 80% and 20% for training and test sets, respectively.

4.2.3. Developed Models for Energy Consumption Predictions

A total of 12 models, including deep learning models such as LSTM, CNN, GRU, and hybrid models (including CNN-LSTM, CNN-GRU, and CNN-bidirectional LSTM), were built and evaluated for their ability to accurately predict energy consumption in smart buildings. The experiments were conducted on minute-by-minute, hourly, and daily time intervals, and the model performance was measured using the root mean square error (RMSE), mean absolute error (MAE), and mean squared error (MSE) metrics.

The proposed CNN-LSTM model was compared to other models in terms of its ability to make short-term, medium-term, and long-term predictions. The dataset was aggregated from minute-by-minute timestamps to hourly and daily timestamps to evaluate the model performance at different time intervals. A sliding-window algorithm [46] with a window size of two was utilized in the experiment, where the model was fed with two consecutive time steps as the inputs and used to predict the next value.

4.3. Evaluation Metrics

The performance of the proposed CNN-LSTM model is evaluated by MAE, MSE, and RMSE. These performance metrics evaluate the variance between the actual and predicted values.

$$MSE = \frac{1}{N}\sum_{i=1}^{n}\left(Y_I - \hat{Y}_I\right)^2 \tag{15}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i - \hat{Y}_I\right| \tag{16}$$

$$RSME = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(Y_I - \hat{Y}_I\right)^2} \tag{17}$$

where Y_I and $\hat{Y}_I$ represent the actual time–series value and the predicted value, respectively. MSE (Equation (15)) measures the average square of the difference between predicted and actual values, MAE (Equation (16)) computes the mean absolute difference between actual and predicted values, and RMSE (Equation (17)) measures the difference between predicted and actual values, it is the square root of MSE [36,37].

4.4. Model Evaluation

To evaluate the developed models, several experiments were conducted. The energy consumption prediction experiment was categorized into minute-by-minute, hourly, and daily predictions. The experiment was carried out by aggregating the data by minutes, hours, and days for the time resolution. The first part of these experiments is meant to evaluate the performance of the hybrid models (CNN-LSTM, CNN_GRU, and CNN-BiLSTM) and compare their performance for minute-by-minute, hourly, and daily forecasting of energy consumption. The second part of the experiment was conducted to evaluate the performance of the hybrid models against the performance of the individual deep neural network models (the LSTM, CNN, and GRU deep learning models). The results were evaluated using RMSE, MSE, and MAE as the performance metrics, as defined in Section 4.3.

4.4.1. Performance Evaluation of Hybrid Models for Minute-by-Minute Time Resolution

The results of the hybrid models for time–series energy consumption prediction are presented in Table 2. The performance of the models was evaluated using the RMSE, MSE, and MAE metrics. The results, as shown in Table 2, demonstrate that the CNN-LSTM hybrid model generally outperformed the other models with an MSE of 0.109, demonstrating its outstanding capability in predicting short-term energy consumption. The experiment confirms that the CNN-LSTM model has a better performance compared to other hybrid methods for short-term energy consumption forecasting.

Table 2. Evaluation results for hybrid learning models for minute-by-minute predictions.

Models	RMSE	MAE	MSE
CNN-LSTM	0.330	0.117	0.109
CNN-GRU	0.369	0.189	0.136
CNN–Bidirectional LSTM	0.3477	0.1339	0.1209

In addition, Figures 9–11 show how the models perform regarding capturing the trends of minute-by-minute energy consumption.

Figure 9 for example, illustrate the CNN-LSTM model showing the predicted trends against the actual trends for minute-by-minute energy consumption. In the figure, the CNN-LSTM model shows a strong alignment between the predicted and actual values of energy consumption. The model captures the small fluctuations in energy consumption, particularly during peak periods of energy consumption. The predicted trends follow the actual trends, with minimal errors, indicating the model's ability to track minute-level energy consumption trends.

Figure 10 represents the CNN-GRU model, which also captures the general trend of the actual energy consumption data, but there are more noticeable errors compared to CNN-LSTM. Whilst the overall pattern of high and low energy consumption is captured, there are instances where the predicted values show some difference from the actual values, especially during periods of high consumption.

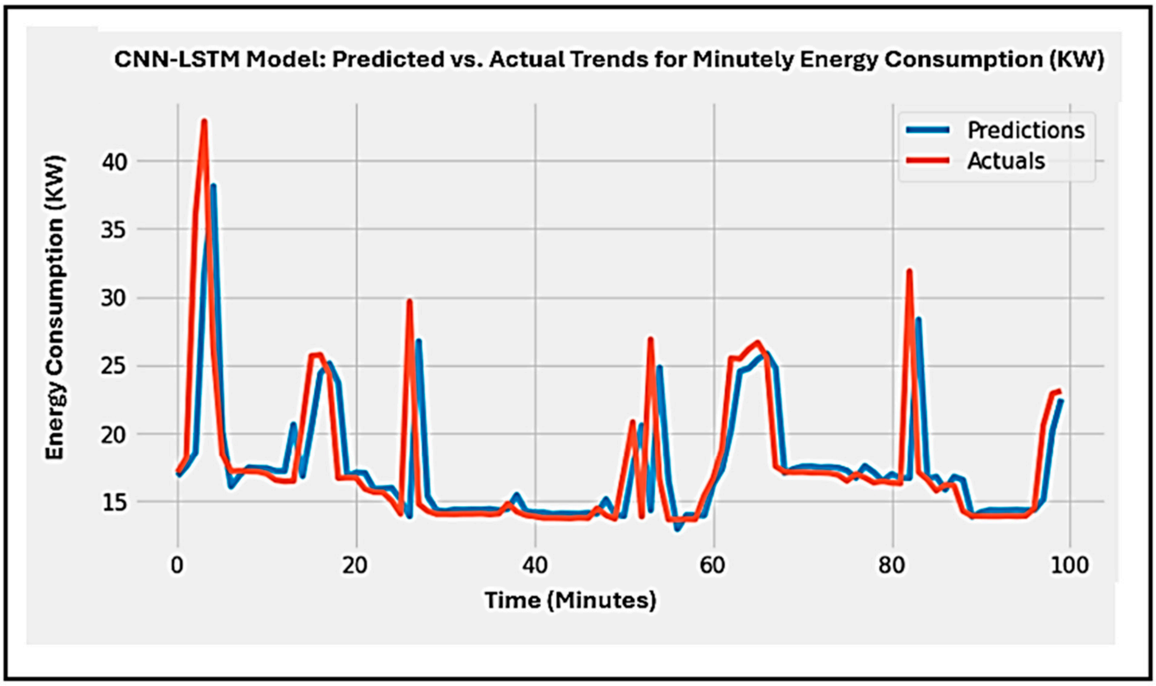

Figure 9. The CNN-LSTM model showing the predicted trends against the actual trends for minute-by-minute energy consumption. In the figure, the CNN-LSTM model shows a strong alignment between the predicted and actual values of energy consumption. The model captures the small fluctuations in energy consumption, particularly during periods of rapid change. The predicted trends follow the actual trends, with minimal deviations, indicating the model's ability to track minute-level energy consumption trends.

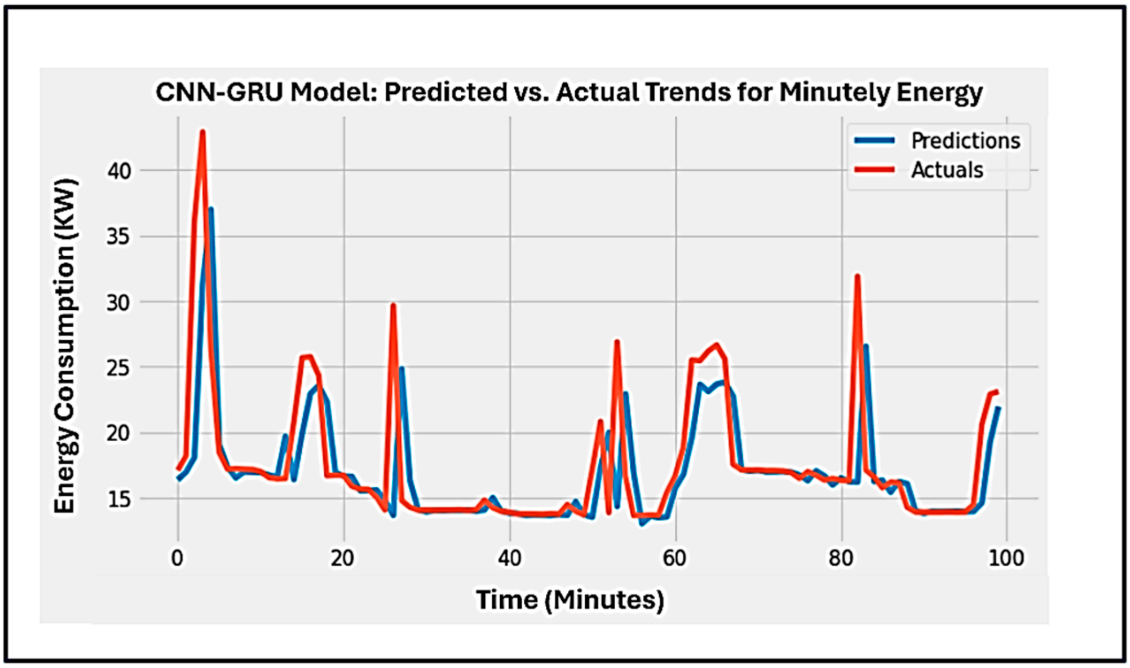

Figure 10. CNN-GRU model for minute-by-minute prediction. The model also captures the general trend of the actual energy consumption data, but there are more noticeable errors compared to CNN-LSTM. Whilst the overall pattern of high and low energy consumption is captured, there are instances where the predicted values exhibit some differences between the actual values, especially during peak consumption periods.

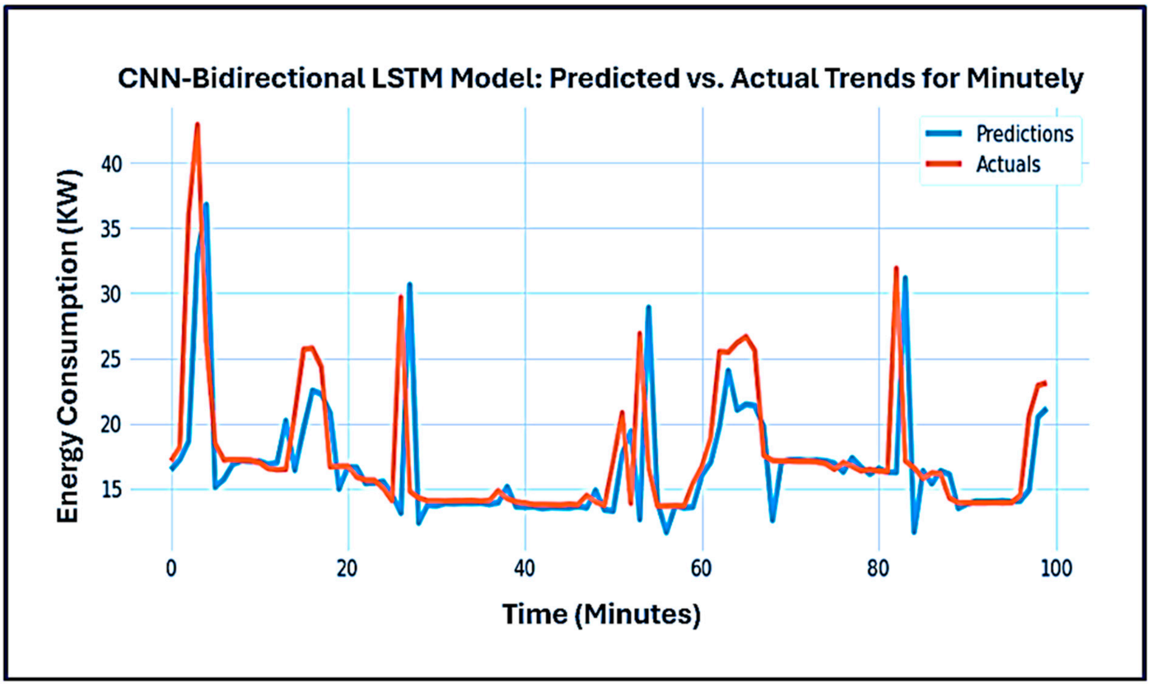

Figure 11. The CNN-BiLSTM model showing the predicted trends against the actual trends for minute-by-minute energy consumption. The CNN-BiLSTM model captures the trends in energy consumption, but it shows more deviations compared to the other two models. The predicted line follows the general direction of the actual data, but at certain points, there are larger gaps between the predicted and actual values, suggesting that this model has more difficulty with minute-level predictions.

.

Then, Figure 11 shows the CNN-BiLSTM model. The CNN-BiLSTM model captures the trends in energy consumption, but it shows more deviations compared to the other two models. The predicted line follows the general direction of the actual data, but at certain points, there are larger gaps between the predicted and actual values, suggesting that this model has more difficulty with minute-level fluctuations compared to other models.

4.4.2. Performance Evaluation of Hybrid Models for Hourly Time Resolution

The evaluation of the proposed models' performances for hourly energy consumption prediction was carried out in the second experiment. The dataset was transformed from minute-by-minute timestamps to hourly time intervals, resulting in a decrease in the number of observations from 790,558 to 13,173. In comparison to other hybrid deep learning models, the proposed CNN-LSTM model performed better in predicting patterns in the dataset. As seen from the results in Table 3, the proposed model's prediction was closer to the actual values, as evidenced by its lowest mean squared error (MSE) of 2.530. Despite the MSE being greater than 0, due to the small size of the dataset used for training the model, the proposed model still demonstrated a remarkable ability to predict hourly energy consumption. Figures 12–14 further show how these models capture the hourly energy consumption trends.

Table 3. Model evaluation for hourly energy consumption.

Models	RMSE	MAE	MSE
CNN-LSTM	1.590	0.895	2.530
CNN-GRU	1.689	1.070	2.855
CNN–Bidirectional LSTM	1.634	0.982	2.678

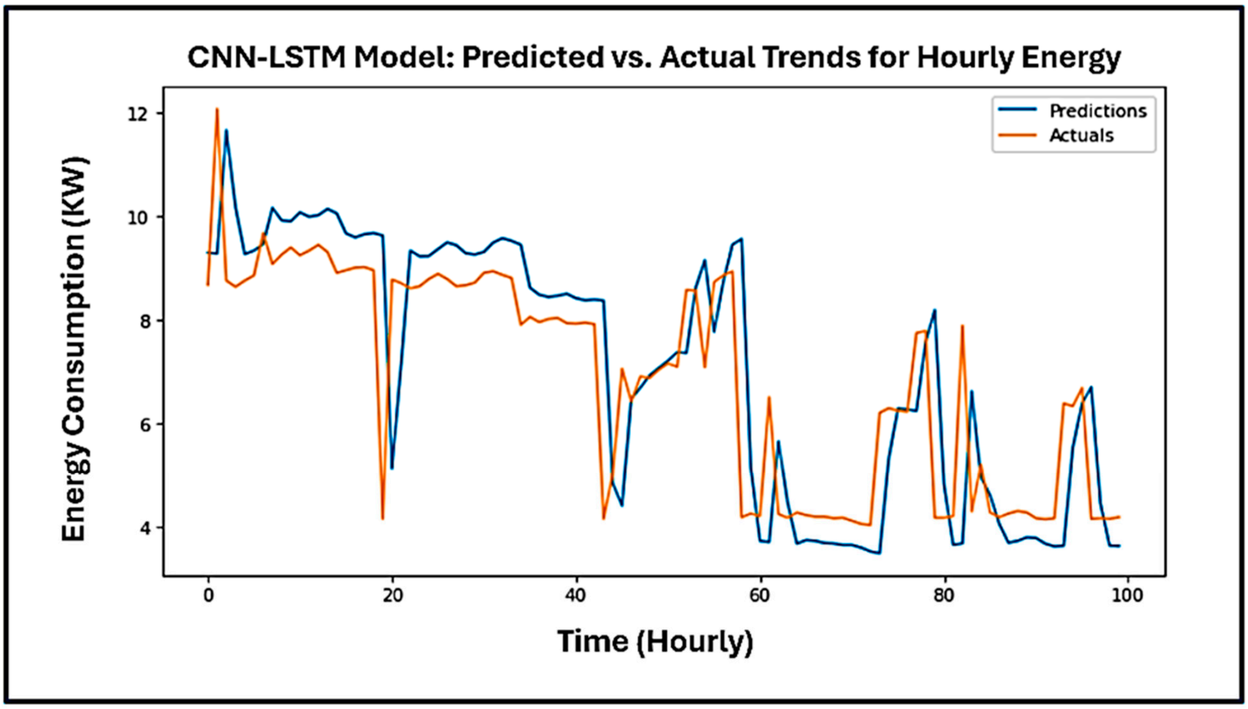

Figure 12. The CNN-LSTM model showing the predicted trends against the actual trends for hourly energy consumption. The CNN-LSTM model captures the broader trends of energy consumption. The predicted line closely follows the actual line, particularly during periods of gradual changes in consumption. However, there are some errors as can be seen at peak period or when consumption declines, where the predicted values are slightly lower than the actual values.

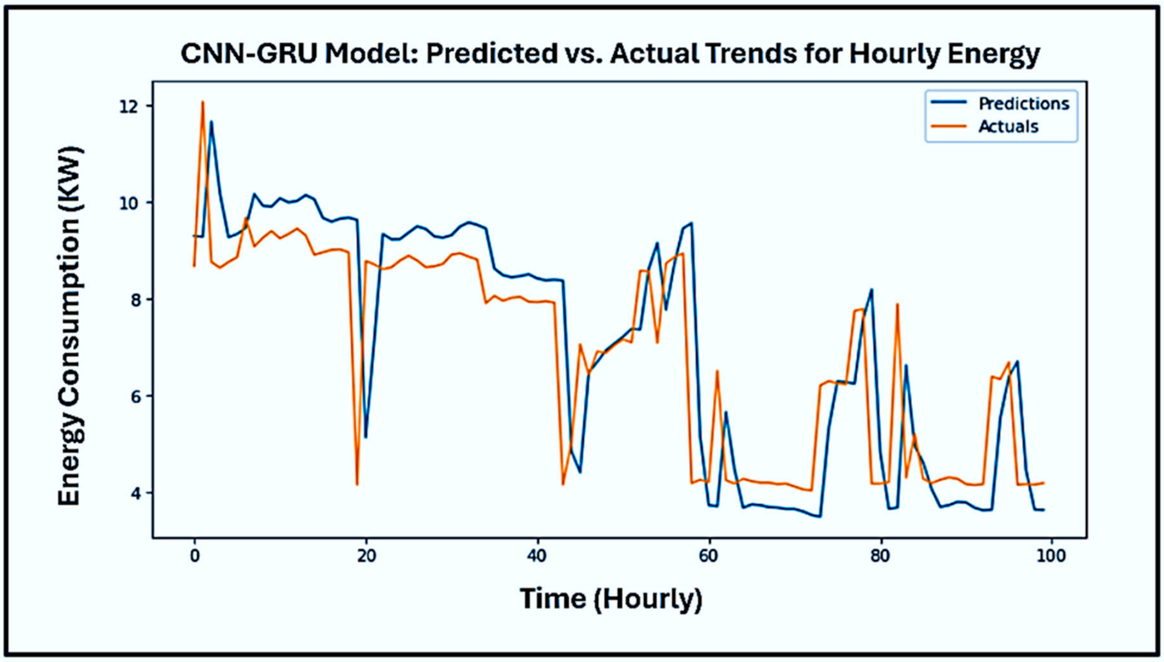

Figure 13. The CNN-GRU model showing predicted trends against the actual trends for hourly energy consumption. The CNN-GRU model's capturing of the hourly trends is like that of the CNN-LSTM model hourly trends, but there are more noticeable deviations. The predicted line and the actual values sometimes overlap, particularly for periods of high and low energy consumption. However, the model still captures the general trends of energy consumption at hourly time resolutions.

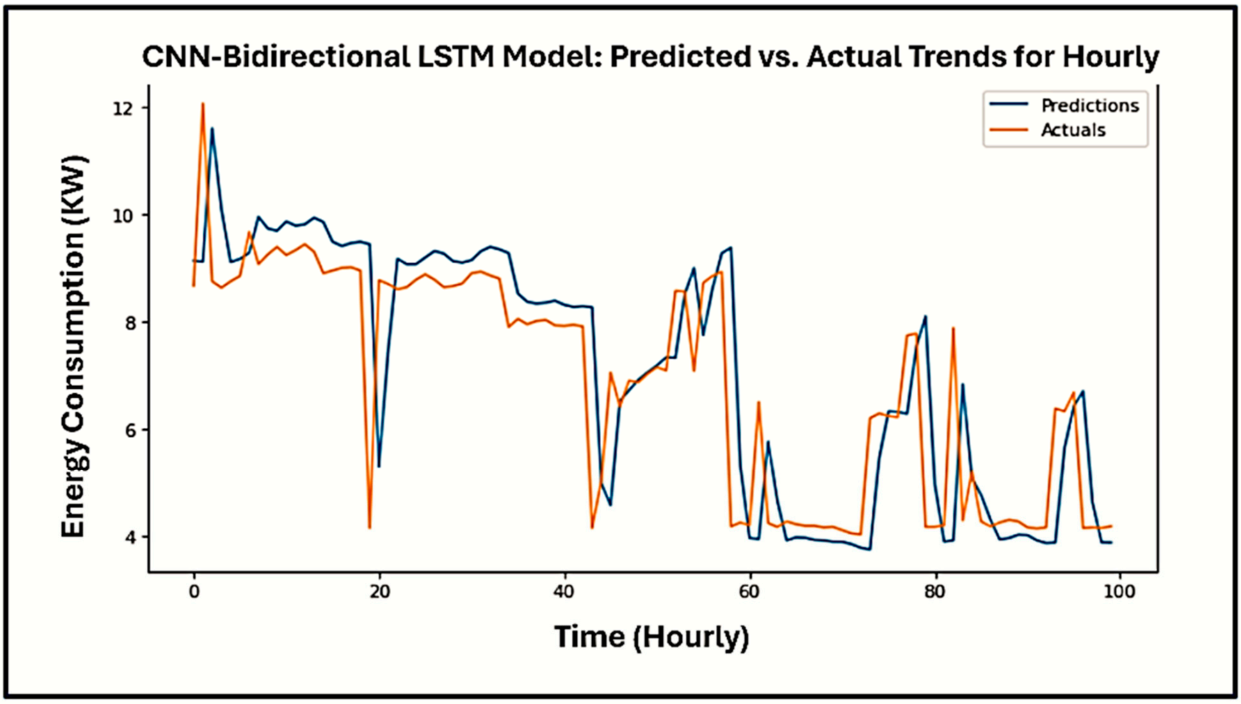

Figure 14. The CNN-BiLSTM model showing predicted trends against the actual trends for hourly energy consumption. This model shows a similar trend, but with larger deviations from the actual trend compared to the CNN-GRU and CNN-LSTM models. This shows that the model is less effective at capturing spikes in hourly consumption.

4.4.3. Performance Evaluation of Hybrid Models for Daily Time Resolution

The third experiment evaluated the performance of the proposed hybrid models for predicting daily energy consumption. The dataset was preprocessed to convert it from hourly timestamps to daily time intervals, resulting in 548 observations. The CNN-LSTM model was compared to other hybrid deep learning models. The results, shown in Table 4, demonstrate that the mean square error values were higher than zero due to the small dataset used in training. Despite this, the CNN-LSTM model showed a better performance (MSE of 28.95) compared to those of CNN-GRU (MSE of 70.04) and CNN-BiLSTM (MSE of 143.4). As can be seen in Figure 15, the CNN-LSTM demonstrates that it is able to forecast energy consumption trends based on daily time resolution. However, Figures 16 and 17 show that both the CNN-GRU and CNN-BiLSTM exhibit poor performances in predicting trends. The poor performances can be attributed to the size of the dataset which is significantly small after the data were preprocessed for the daily time resolution.

Table 4. Model evaluation for daily energy consumption prediction.

Models	RMSE	MAE	MSE
CNN-LSTM	5.380	3.649	28.95
CNN-GRU	8.366	7.270	70.04
CNN–Bidirectional LSTM	11.97	11.22	143.4

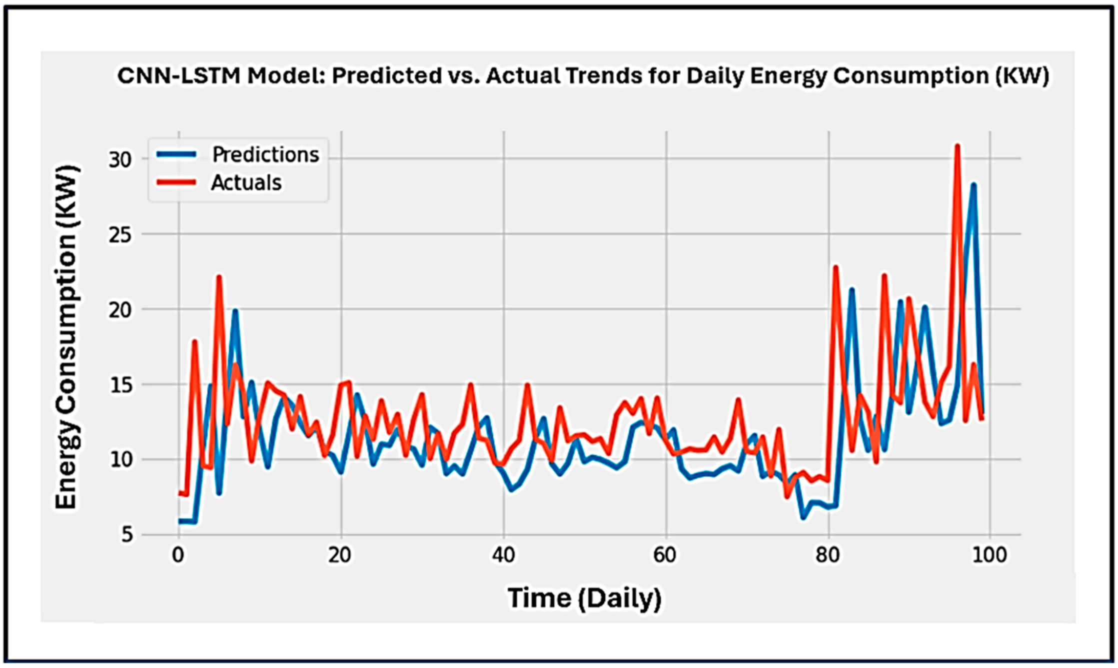

Figure 15. The CNN-LSTM model showing the predicted trends against the actual trends for daily energy consumption. In the daily predictions, the CNN-LSTM model captures the general trend but shows larger errors compared to minute-by-minute and hourly predictions. The predicted trends still follow the direction of the actual trends, but there are slightly larger errors between the predicted and actual trends, particularly during peak periods of consumption.

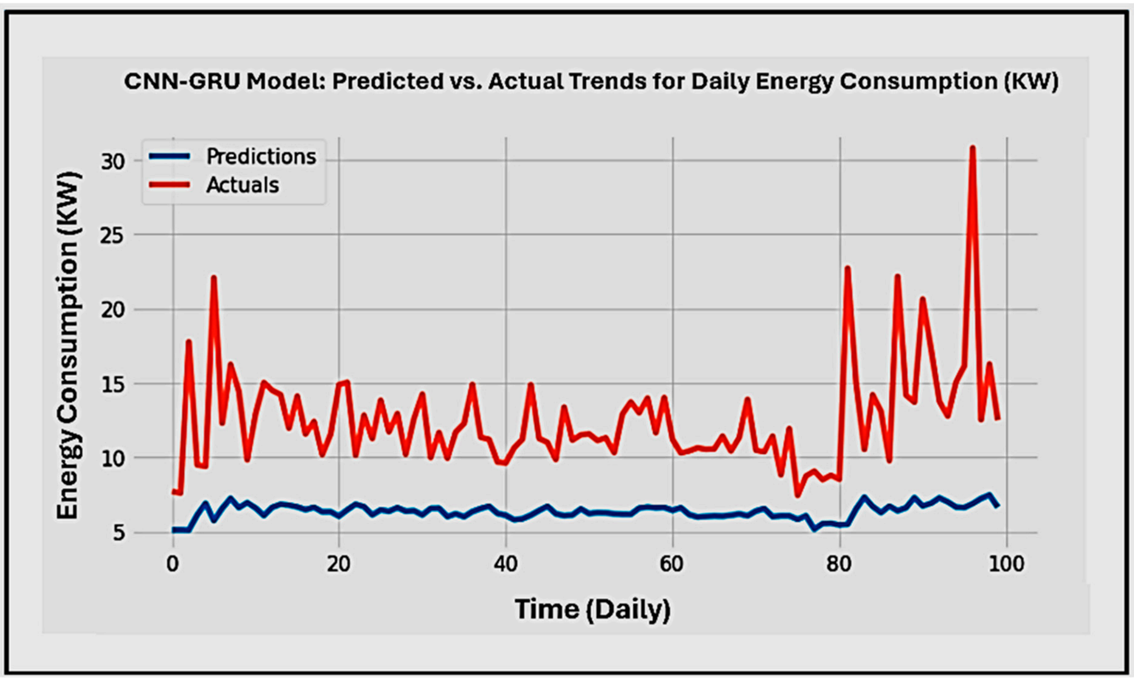

Figure 16. The CNN-GRU model showing the predicted trends against the actual trends for daily energy consumption. As can be seen, unlike the CNN-LSTM model, the CNN-GRU model is unable to capture the actual trends. With an MSE of 70.04, the CNN-GRU model shows a significant increase in error compared to CNN-LSTM. The higher error rates indicate that this model does not perform well in predicting daily energy consumption trends.

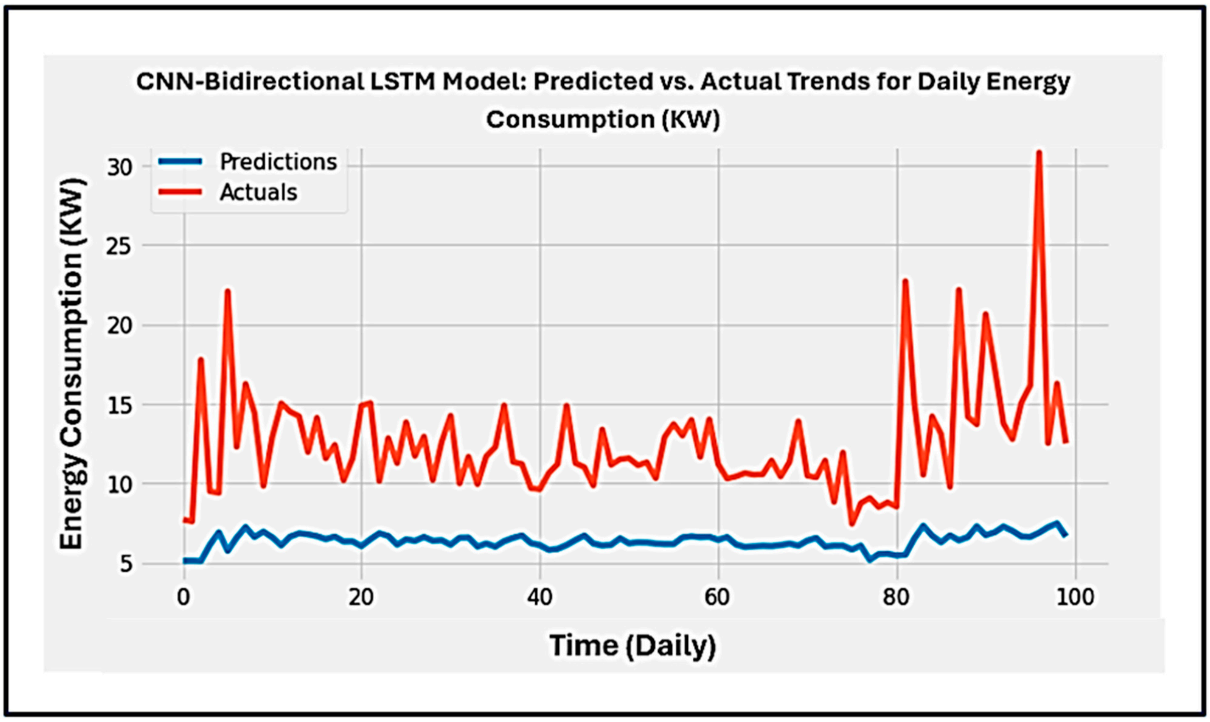

Figure 17. The CNN-BiLSTM model showing predicted trends against the actual trends for hourly energy consumption. The CNN–Bidirectional LSTM model has the worst performance in predicting the daily trends as can be seen.

4.4.4. Performance Evaluation of All Models for Minute-by-Minute, Hourly and Daily Forecasting

In this experiment, the performance of the individual models was conducted to compare their performance of the hybrid models with the individual models. Table 5 shows each model's performance according to changes in time. As the time resolution decreases, the error rate increases. This is because, at each stage of aggregating the dataset, the number of observations keeps decreasing leaving smaller data to be trained by the model. Furthermore, a deep neural network requires a larger amount of data to increase the performance. However, at each stage of the time change, the proposed model outperformed the other models, which proves that the proposed model is superior.

Table 5. Accuracies of both the individual deep learning model and their hybrid counterparts, showing the performances of the models for minute-by-minute, hourly, and daily energy consumptions.

Models	Time Resolution	Error Metrics		
		RMSE	**MAE**	**MSE**
LSTM	Minute-by-minute	0.329	0.120	0.188
CNN		0.385	0.220	0.148
GRU		0.335	0.164	0.112
CNN-LSTM		0.330	0.117	0.109
CNN-GRU	Hourly	1.689	1.070	2.855
CNN-BiLSTM		1.634	0.982	2.678
CNN-LSTM		1.590	0.895	2.530
CNN-GRU	Daily	8.366	7.270	70.04
CNN-BiLSTM		11.97	11.22	143.4
CNN-LSTM		5.380	3.649	28.95

5. Discussion

The goal of the work presented in this article is to evaluate the performance of hybrid deep learning models for predicting energy consumption in smart buildings at various time resolutions. In the experimental validation, a total of 12 deep neural network and hybrid neural network models were built and evaluated. The experiments were conducted to understand the performance of these models for capturing trends in energy consumption at various time intervals as well as to evaluate the accuracy of their predictions. The evaluations examine the performance of models for capturing temporal patterns in energy consumption at minute-by-minute, hourly, and daily intervals. The results show that nearly all models performed well in predicting the trends and patterns present in the dataset.

For minute interval prediction, the CNN-LSTM model has the best forecasting performance in terms of MSE loss error, achieving 0.109. This result demonstrates that the CNN-LSTM hybrid model is the best model amongst those evaluated for predicting short-term energy consumption. Although the forecasting errors of the other models (CNN–Bidirectional LSTM and CNN-GRU) are not significantly higher compared to that of the CNN-LSTM, it performs better when predicting at minute intervals during the peak periods than the other models.

Additionally, in the second set of experiments on hourly forecasting, the results also provide evidence that the CNN-LSTM model performs better than the other models. With a decrease in the number of observations, the proposed model was still able to predict trends and achieved the lowest MSE of 2.530. CNN-BiLSTM followed, with an MSE of 2.678.

Similarly, the third set of experiments was conducted to evaluate the performance of the models for the daily prediction of energy consumption. Although the number of datapoints significantly reduced when the dataset was preprocessed for daily predictions, the results also demonstrate that CNN-LSTM performs better than the other models, achieving an MSE of 28.95.

Generally, the results of this study provide strong evidence of the efficacy of the CNN-LSTM model for energy consumption prediction across multiple time resolutions. The model consistently outperformed both the standard deep learning models (such as LSTM, CNN, and GRU) and the hybrid models (such as CNN-GRU and CNN–Bidirectional LSTM) when evaluated on key performance metrics such as RMSE, MAE, and MSE. Although, CNN-LSTM is more complex and more computationally expensive than the GRU, it can be seen from these results that it has the capability to capture long-term dependencies in sequences of the energy consumption dataset. This performance advantage is attributed to the model's ability to capture both spatial and temporal features as well as having more gates to control the flow of information through the model's networks, proving to be particularly good for peak energy consumption prediction. It was also observed that the CNN–Bidirectional LSTM, which combines two LSTMs in forward and backward directions, failed to achieve the same or better level of performance than the CNN-LSTM models.

However, the study has some limitations; all models, including the CNN-LSTM, exhibited significantly higher error rates when predicting daily energy consumption. This decline in performance is likely due to data aggregation, which reduced the number of observations available for model training. The findings underscore the importance of a large and diverse dataset, particularly when applying deep learning models to time–series forecasting tasks.

Nonetheless, the study provides some valuable insights into the benefits and challenges of applying hybrid deep learning models in the context of energy consumption forecasting, with the CNN-LSTM model proving robust in short- and medium-term resolution predictions.

6. Conclusions and Future Work

In this article, we investigated the performance of various deep learning models such as CNN, LSTM, GRU, and bidirectional LSTM for predicting and forecasting energy consumption in smart buildings. The article also investigated the capacity of these models in combination with the CNN architecture to forecast energy consumption trends for various time resolutions such as minutes, hours, and days. The results show that CNN-LSTM model produced the best prediction accuracy among the three hybrid models evaluated for forecasting energy usage trends over various time resolutions. The CNN-LSTM model demonstrates a better performance for extracting complex latent spatial and temporal features from historical energy consumption data, enabling it to predict consumption trends better than other individual models, including the LSTM, CNN, and GRU. The model particularly performed better for short time resolutions, such as minute-by-minute predictions, achieving the lowest MSE of 0.109 for minute-level forecasts; this illustrated its efficacy in forecasting complex fluctuations in energy usage.

The results also show that, for hourly forecasting, the CNN-LSTM hybrid model achieved an MSE of 2.530, the lowest among other evaluated models, such as the CNN-GRU and the CNN–Bidirectional LSTM. Additionally, the CNN-LSTM hybrid model demonstrates a better performance for peak load predictions and trends forecasting, consistently outperforming individual deep learning and other hybrid models.

Despite its better performance, the CNN-LSTM demonstrated a poor performance when predicting long-term energy consumption. In the future, we would like to investigate how additional data such as activities, behaviours, and contextual information of a smart building's occupants would impact the prediction and forecasting of energy usage. In addition, we would like to investigate other deep-learning-based models, such as large language models (LLMs), for energy consumption prediction and forecasting.

Author Contributions: Conceptualization, A.O. and F.I.; Data curation, F.I.; Formal analysis, A.O., F.I., J.E.A. and A.I.; Investigation, A.O., F.I., J.E.A. and A.I.; Methodology, A.O. and F.I.; Project administration, A.O. and F.I.; Resources, F.I.; Software, F.I.; Supervision, A.O.; Validation, A.O. and F.I.; Visualization, A.O., F.I., J.E.A. and A.I.; Writing—original draft, A.O. and F.I.; Writing—review and editing, A.O., F.I., J.E.A. and A.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Source of dataset [45].

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

References

1. Olu-Ajayi, R.; Alaka, H.; Sulaimon, I.; Sunmola, F.; Ajayi, S. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. *J. Build. Eng.* **2022**, *45*, 103406. [CrossRef]
2. Ozili, P.K. Sustainability accounting. *SSRN Electron. J.* **2021**. [CrossRef]
3. Pérez-Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [CrossRef]
4. Hu, S.; Hoare, C.; Raftery, P.; O'Donnell, J. Environmental and energy performance assessment of buildings using scenario modelling and fuzzy analytic network process. *Appl. Energy* **2019**, *255*, 113788. [CrossRef]
5. Sharma, S.; Xu, Y.; Verma, A.; Panigrahi, B.K. Time-Coordinated Multienergy Management of Smart Buildings Under Uncertainties. *IEEE Trans. Ind. Inform.* **2019**, *15*, 4788–4798. [CrossRef]
6. Gupta, A.; Badr, Y.; Negahban, A.; Qiu, R.G. Energy-efficient heating control for smart buildings with deep reinforcement learning. *J. Build. Eng.* **2021**, *34*, 101739. [CrossRef]
7. Bourhnane, S.; Abid, M.R.; Lghoul, R.; Zine-Dine, K.; Elkamoun, N.; Benhaddou, D. Machine learning for energy consumption prediction and scheduling in smart buildings. *SN Appl. Sci.* **2020**, *2*, 297. [CrossRef]
8. Xu, D.; Yang, S.; Zhang, H.; Xu, Q.; Li, Z.; Lu, Z.; Chen, W. A Classified Identification Deep-Belief Network for Predicting Electric-Power Load; A Classified Identification Deep-Belief Network for Predicting Electric-Power Load. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20–22 October 2018.
9. Dai, X.; Liu, J. and Zhang, X. A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings. *Energy Build.* **2020**, *223*, 110159. [CrossRef]
10. Zhao, C.; Zhang, M.; Wang, W. Exploring the influence of severe haze pollution on residents' intention to purchase energy-saving appliances. *J. Clean. Prod.* **2019**, *212*, 1536–1543. [CrossRef]
11. Mariano-Hernández, D.; Hernández-Callejo, L.; Solís, M.; Zorita-Lamadrid, A.; Duque-Perez, O.; Gonzalez-Morales, L.; Santos-García, F. A Data-Driven Forecasting Strategy to Predict Continuous Hourly Energy Demand in Smart Buildings. *Appl. Sci.* **2021**, *11*, 7886. [CrossRef]

12. Afzal, S.; Ziapour, B.M.; Shokri, A.; Shakibi, H.; Sobhani, B. Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms. *Energy* **2023**, *282*, 128446. [CrossRef]
13. Sekhar, C.; Dahiya, R. Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand. *Energy* **2023**, *268*, 126660. [CrossRef]
14. Cai, W.; Wen, X.; Li, C.; Shao, J.; Xu, J. Predicting the energy consumption in buildings using the optimized support vector regression model. *Energy* **2023**, *273*, 127188. [CrossRef]
15. Mubarak, H.; Stegen, S.; Bai, F.; Abdellatif, A.; Sanjari, M.J. Enhancing interpretability in power management: A time-encoded household energy forecasting using hybrid deep learning model. *Energy Convers. Manag.* **2024**, *315*, 118795. [CrossRef]
16. Sunder, R.; Paul, V.; Punia, S.K.; Konduri, B.; Nabilal, K.V.; Lilhore, U.K.; Lohani, T.K.; Ghith, E.; Tlija, M. An advanced hybrid deep learning model for accurate energy load prediction in smart building. *Energy Explor. Exploit.* **2024**, *42*, 01445987241267822. [CrossRef]
17. El Maghraoui, A.; Ledmaoui, Y.; Laayati, O.; El Hadraoui, H.; Chebak, A. Smart Energy Management: A Comparative Study of Energy Consumption Forecasting Algorithms for an Experimental Open-Pit Mine. *Energies* **2022**, *15*, 4569. [CrossRef]
18. Ma, H. Prediction of industrial power consumption in Jiangsu Province by regression model of time variable. *Energy* **2022**, *239*, 122093. [CrossRef]
19. Wang, M.; Wang, W.; Wu, L. Application of a new grey multivariate forecasting model in the forecasting of energy consumption in 7 regions of China. *Energy* **2022**, *243*, 123024. [CrossRef]
20. Gellert, A.; Fiore, U.; Florea, A.; Chis, R.; Palmieri, F. Forecasting Electricity Consumption and Production in Smart Homes through Statistical Methods. *Sustain. Cities Soc.* **2022**, *76*, 103426. [CrossRef]
21. Zhang, X.M.; Grollinger, K.; Capretz, M.A.; Seewald, L. Forecasting Residential Energy Consumption: Single Household Persepctive. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications, ICMLA, Orlando, FL, USA, 17–20 December 2018; pp. 110–117.
22. Catalina, T.; Iordache, V.; Caracaleanu, B. Multiple regression model for fast prediction of the heating energy demand. *Energy Build.* **2013**, *57*, 302–312. [CrossRef]
23. Pombeiro, H.; Santos, R.; Carreira, P.; Silva, C.; Sousa, J.M.C. Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. *Energy Build.* **2017**, *146*, 141–151. [CrossRef]
24. Shao, X.; Kim, C.-S.; Sontakke, P. Accurate Deep Model for Electricity Consumption Forecasting Using Multi-Channel and Multi-Scale Feature Fusion CNN–LSTM. *Energies* **2020**, *13*, 1881. [CrossRef]
25. Azuana Ramli, N.; Laili Mazlan, N.; Jamilatul Awalin, L.; Badrulhisham Ismail, M.; Kassim, A.; Menon, A. A Smart Building Energy Management Using Internet of Things (IoT) and Machine Learning. 2020. Available online: https://www.researchgate.net/publication/342956749 (accessed on 23 February 2023).
26. Shapi, M.K.M.; Ramli, N.A.; Awalin, L.J. Energy consumption prediction by using machine learning for smart building: Case study in Malaysia. *Dev. Built Environ.* **2021**, *5*, 100037. [CrossRef]
27. Wang, Y.; Xia, Q.; Kang, C. Secondary forecasting based on deviation analysis for short-term load forecasting. *IEEE Trans. Power Syst.* **2011**, *26*, 500–507. [CrossRef]
28. Cervantes, J.; Li, X.; Yu, W.; Li, K. Support vector machine classification for large data sets via minimum enclosing ball clustering. *Neurocomputing* **2008**, *71*, 611–619. [CrossRef]
29. Kiprijanovska, I.; Stankoski, S.; Ilievski, I.; Jovanovski, S.; Gams, M.; Gjoreski, H. Houseec: Day-ahead household electrical energy consumption forecasting using deep learning. *Energies* **2020**, *13*, 2672. [CrossRef]
30. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6. [CrossRef]
31. Morcillo-Jimenez, R.; Mesa, J.; Gómez-Romero, J.; Vila, M.A.; Martin-Bautista, M.J. Deep learning for prediction of energy consumption: An applied use case in an office building. *Appl. Intell.* **2024**, *54*, 5813–5825. [CrossRef]
32. Palak, M.; Revati, G.; Sheikh, A. Smart Building Energy Management using Deep Learning Based Predictions. In Proceedings of the 2021 North American Power Symposium, College Station, TX, USA, 14–16 November 2021. [CrossRef]
33. Sajjad, M.; Khan, Z.A.; Ullah, A.; Hussain, T.; Ullah, W.; Lee, M.Y.; Baik, S.W. A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting. *IEEE Access* **2020**, *8*, 143759–143768. [CrossRef]
34. Graves, A.; Jaitly, N.; Mohamed, A.R. Hybrid Speech Recognition with Deep Bidirectional LSTM. ieeexplore.ieee.org. Available online: https://ieeexplore.ieee.org/abstract/document/6707742/ (accessed on 20 March 2023).
35. Yan, K.; Li, W.; Ji, Z.; Qi, M.; Du, Y. A Hybrid LSTM Neural Network for Energy Consumption Forecasting of Individual Households. *IEEE Access* **2019**, *7*, 157633–157642. [CrossRef]
36. Kim, T.Y.; Cho, S.B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* **2019**, *182*, 72–81. [CrossRef]
37. Khalil, M.; McGough, A.S.; Pourmirza, Z.; Pazhoohesh, M.; Walker, S. Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption—A systematic review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105287. [CrossRef]
38. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* **2020**, *53*, 5929–5955. [CrossRef]

39. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [CrossRef] [PubMed]
40. Beck, M.; Pöppel, K.; Spanring, M.; Auer, A.; Prudnikova, O.; Kopp, M.; Klambauer, G.; Brandstetter, J.; Hochreiter, S. xlstm: Extended long short-term memory. *arXiv* **2024**, arXiv:2405.04517.
41. Shah, F.; Naik, T.; Vyas, N. LSTM based music generation. In Proceedings of the 2019 International Conference on Machine Learning and Data Engineering (iCMLDE), Taipei, Taiwan, 2–4 December 2019; pp. 48–53. [CrossRef]
42. Ookura, S.; Mori, H. An efficient method for wind power generation forecasting by LSTM in consideration of overfitting prevention. *IFAC-Pap.* **2020**, *53*, 12169–12174. [CrossRef]
43. Sherstinsky, A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [CrossRef]
44. Cahuantzi, R.; Chen, X.; Güttel, S. A comparison of LSTM and GRU networks for learning symbolic sequences. *arXiv* **2021**, arXiv:2107.02248.
45. Pipattanasomporn, M.; Chitalia, G.; Songsiri, J.; Aswakul, C.; Pora, W.; Suwankawin, S.; Audomvongseree, K.; Hoonchareon, N. CU-BEMS, smart building electricity consumption and indoor environmental sensor datasets. *Sci. Data* **2020**, *7*, 241. [CrossRef]
46. Figo, D.; Diniz, P.C.; Ferreira, D.R.; Cardoso, J.M.P. Pre-processing techniques for context recognition from accelerometer data. *Pers. Ubiquitous Comput.* **2010**, *14*, 645–662. [CrossRef]