

Article

Forecasting Building Energy Consumption Using Ensemble Empirical Mode Decomposition, Wavelet Transformation, and Long Short-Term Memory Algorithms

Shuo-Yan Chou ^{1,2}, Anindhita Dewabharata ², Ferani E. Zulvia ^{3,*}  and Mochamad Fadil ³

¹ Taiwan Building Technology Center, National Taiwan University of Science and Technology, Taipei 106, Taiwan; sychou@mail.ntust.edu.tw

² Department of Industrial Management, National Taiwan University of Science and Technology, Taipei 106, Taiwan; d10101801@mail.ntust.edu.tw

³ Department of Logistics Engineering, Universitas Pertamina, Jakarta 12220, Indonesia; mochamadfadildil@gmail.com

* Correspondence: feranieva.z@universitaspertamina.ac.id; Tel.: +62-21-290-44308

Abstract: A building, a central location of human activities, is equipped with many devices that consume a lot of electricity. Therefore, predicting the energy consumption of a building is essential because it helps the building management to make better energy management policies. Thus, predicting energy consumption of a building is very important, and this study proposes a forecasting framework for energy consumption of a building. The proposed framework combines a decomposition method with a forecasting algorithm. This study applies two decomposition algorithms, namely the empirical mode decomposition and wavelet transformation. Furthermore, it applies the long short term memory algorithm to predict energy consumption. This study applies the proposed framework to predict the energy consumption of 20 buildings. The buildings are located in different time zones and have different functionalities. The experiment results reveal that the best forecasting algorithm applies the long short term memory algorithm with the empirical mode decomposition. In addition to the proposed framework, this research also provides the recommendation of the forecasting model for each building. The result of this study could enrich the study about the building energy forecasting approach. The proposed framework also can be applied to the real case of electricity consumption.

Keywords: energy building; LSTM; decomposition; empirical mode decomposition; wavelet transformation



Citation: Chou, S.-Y.; Dewabharata, A.; Zulvia, F.E.; Fadil, M. Forecasting Building Energy Consumption Using Ensemble Empirical Mode Decomposition, Wavelet Transformation, and Long Short-Term Memory Algorithms. *Energies* **2022**, *15*, 1035. <https://doi.org/10.3390/en15031035>

Academic Editor: Fabrizio Ascione

Received: 21 November 2021

Accepted: 6 January 2022

Published: 29 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Energy is one of the basic needs of human life. In the last few decades, energy consumption has increased significantly due to several factors, such as a population increase, indoor human activities, an increasing number of buildings, and global climate change. As most human activities are conducted in buildings, it consumes most of the electricity. In the US and the European Union, 40% of energy consumption comes from buildings [1].

Buildings can be categorized into several types based on their function: industry, transportation, housing or residential, commercial, public services, agriculture, fisheries, and other sectors. Among all of these types, housing/residential buildings such as housing complexes and apartments consume up to 30% of the total energy buildings worldwide [2]. In some countries, the energy consumption of commercial buildings such as offices, hospitals, shopping centers, restaurants, warehouses, and others is relatively high because they use many devices continuously [3]. The EIA [4] predicts that energy consumption in buildings will grow by an average of 1.3% per year from 2018 to 2050.

Although energy consumption is predicted to increase continuously, more accurate energy demand forecasting is essential. An accurate energy demand forecast will be essential information to make a decision in many different fields, including scheduling,

operations, monitoring, and others [5]. Forecasting has an important role in different perspectives for future energy development regionally, nationally, and globally [6]. From the perspective of city stakeholders, it is helpful to understand energy consumption and its variations across types of urban buildings and their user profiles. Energy allocation based on an accurate prediction can improve building policy [7]. Accurate energy consumption forecasting can also avoid overestimating and underestimating resource allocation [8]. From the building owner's or operator's point of view, energy forecasting is essential for designing daily operational rules for the building [9]. For the power company and government, city-scale energy consumption predictions help to optimize the electricity distribution and production scenario [9]. Therefore, energy management with accurate forecasting is needed because accuracy allows management to allocate energy needs due to the significant demand to avoid unexpected power outages and reduce operational costs [10]. The primary purpose of this study is to propose a forecasting framework for energy consumption.

Forecasting itself can be classified based on the timespan; short-term, mid-term, and long-term [11]. For example, short-term forecasting with a time span of hourly data for one week [12] will have higher accuracy than mid-term and long-term forecasting. It helps the building management in operation [13]. It also can improve the automation of the energy system of a building [14] at the management level.

Many studies accommodate this problem to improve an accurate energy analysis system, one of which is a research work that forecasts electrical energy consumption in buildings. However, several factors become challenges for researchers in making accurate forecasts [5]. The challenging factors include the limitation of the physical method which requires high computational complexity. On the other hand, the statistics methods cannot perform well for nonlinear hidden feature data.

The proposed forecasting framework combines a time series decomposition method with recurrent neural network algorithms to obtain an accurate forecasting model. There are two decomposition algorithms utilized in this study. They are the ensemble empirical mode decomposition (EEMD) and wavelet transformation (WT). Some previous studies have applied forecasting energy demand using the long short-term memory (LSTM) algorithm [15–17]. Peng, Wang, Xia, and Gao [15] and Somu, MR, and Ramamritham [17] applied the LSTM with WT. On the other hand, Gao, Ruan, and Buildings [16] enhanced the LSTM with a feature attention process. These studies show that the LSTM algorithm is a promising algorithm for energy consumption prediction. However, further improvements should be made to get a better result with a high generalization. Thus, this study proposes a forecasting framework, which consists of data preprocessing, data decomposition using EEMD and WT, and the aggregation process for the final prediction.

This study evaluates the proposed forecasting framework's performance using six benchmark datasets and a real study case of electricity consumption of a university building in Taiwan. Each dataset consists of several buildings. Thus, a total of 20 buildings are evaluated in this study.

2. Literature Review

This section reviews some fundamental theories applied in this research. It includes time series forecasting, decomposition algorithms, and recurrent neural network.

2.1. Time Series Forecasting

Forecasting is the process of predicting future events [11]. Runge and Zmeureanu [18] define forecasting as the process of estimating a value derived from current values or historical data to determine the value that will be generated in the future. In forecasting energy consumption, the data used are time-series data. Electrical energy consumption data are time-series data observed based on the same time interval [19]. Therefore, forecasting electrical energy consumption is categorized as time-series data forecasting. Its goal is to predict events that will occur in the future based on a series of events or history [5].

Problems in forecasting are generally packaged in the form of a time period, such as days, weeks, or months [11]. Based on the time horizon, forecasting is divided into several categories [20], those are:

1. Short-term

The coverage period of this forecast is hours, days, or weeks ahead. This type of forecasting manages daily operational activities, purchase planning, and evaluation [21].

2. Medium-term

The scope of the forecasting period is 1–2 years in the future [11]. This forecasting category is very useful for the strategic planning of demand [21].

3. Long-term

The coverage period of this forecast is about 1–50 years in the future. It is commonly used in a planning system or energy plant installation [21].

Many techniques have been used for forecasting data on electricity consumption in buildings. To obtain an accurate forecast, the methodology should consider data complexity, including weather conditions, lighting systems, and other systems that involve the use of electricity [22]. According to Deb, Zhang, Yang, Lee, and Shah [22], there are two categories of forecasting techniques: the data-driven technique and the deterministic technique. Table 1 shows several advantages and disadvantages of both techniques.

Table 1. The Advantages and Disadvantages of Data-Driven and Deterministic Technique.

Techniques	Advantages	Disadvantages
Data-driven	Faster calculations with real-time data	Historical data are needed
	Can be applied for non-linear problems	Solution for a case may not be suitable for other cases
Deterministic	Refers to the physical science of the building	It is difficult to create scenarios that refer to the actual model
	Does not require historical data	Requires detail building properties data
	One solution can be applied to many cases	-

Source: [22].

Bourdeau, et al. [23] divide energy forecasting methods in buildings into three categories, namely physical (white box), data-driven (black box), and hybrid (gray box). The data-driven (black box) method uses machine learning and statistics to build and model energy in buildings [23]. The physical method (white box) can be interpreted the same as the deterministic method above. In contrast, the hybrid method combines knowledge and information from physical and data-driven methods. Most of the methods used by researchers are data-driven (black box).

2.2. Recurrent Neural Network

Recurrent neural network (RNN) is a very popular model used to solve forecasting time series because it can retrieve information obtained from time-series data [24]. The RNN develops the forecasting model based on the input in the current state and inputs from the previous states [25]. RNN extends an artificial neural network (ANN) that connects neurons in the same hidden layer [26] the structure as illustrated in Figure 1.

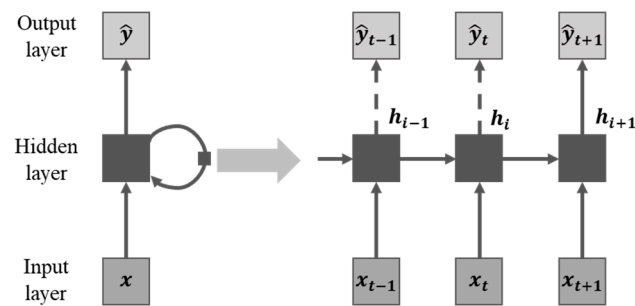


Figure 1. Structure of recurrent neural network [26].

The RNN structure consists of a loop that can store information from the past [27]. Although the RNN has good forecasting capabilities, it cannot handle data dimensions of long-term dependencies [24]. The long short-term memory (LSTM) improves RNN, overcoming the long-term dependencies problems [28]. The LSTM solves the vanishing gradient problem due to not overcoming data with long-term dependencies [24]. The LSTM uses a component called memory blocked [28]. The vanishing gradient problem can be eliminated by applying the gate mechanism and memory cells that replace the nodes in it [29] as illustrated in Figure 2.

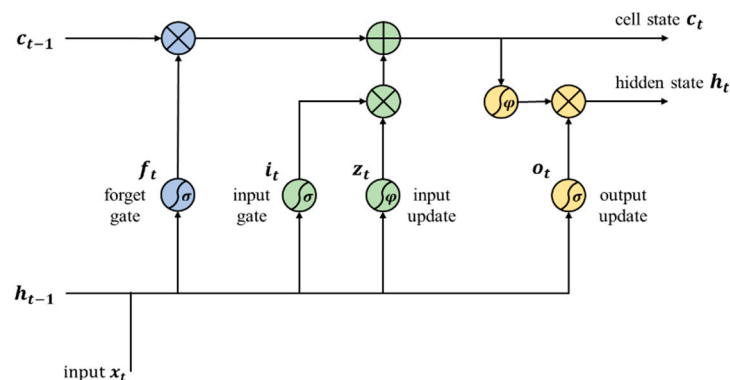


Figure 2. Structure of long short-term memory (LSTM) [26].

Memory blocked is a subnet containing repeatedly connected functional modules called memory cells and gates [26]. Memory cells are tasked with remembering the state of the network and the gates formed to regulate information patterns [26]. LSTM is divided into three parts, namely input gates, output gates, and forget gates. Input gates function to control the information entering memory cells, forget gates function to remember the amount of information passing through the network, and output gates function to control the amount of information used to perform calculations [26].

2.3. Ensemble Empirical Mode Decomposition

The empirical mode decomposition (EMD) is a pre-processing method used for non-stationary data [30], such as energy-time series data in buildings. Non-stationary data will be decomposed into several components named intrinsic mode functions (IMF) as illustrated in Figure 3 [31]. The EMD method shows its superiority in time series data forecasting [30]. Thus, the ensemble empirical mode decomposition (EEMD) was proposed to improve the EMD [32]. This method was created to overcome the mixing mode problems in EMD [32,33]. It also aims to extract the existing oscillation function [30]. Mode mixing is a problem that arises because one IMF component has a different signal scale or the same signal scale in different IMFs [34]. To bypass the mixing mode, EEMD follows the EMD steps on the original time series $x(t)$, ($t = 1, 2, \dots, t$), to get a set of IMF added by Gaussian white noises. Then, the average from the IMF applies as the final decomposition result [30].

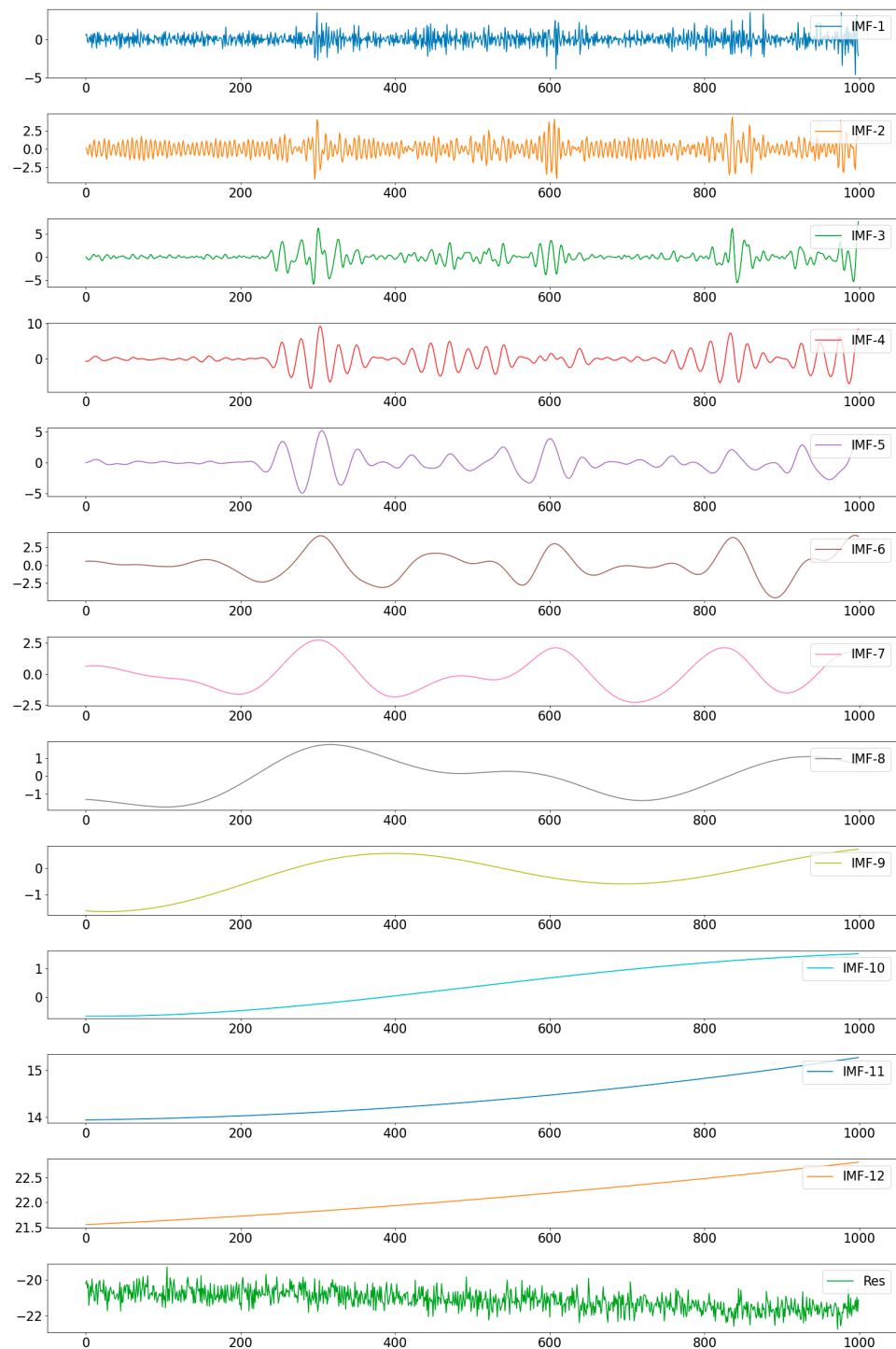


Figure 3. Intrinsic mode function (IMF) [5].

The steps of EEMD as explained below [30]:

1. Add the white noise into the original data $x(t)$ to get the new construction:

$$x^i(t) = x(t) + w^i(t), \quad (1)$$

2. Describes the time series which have been added the white noise into n th IMF $c_j^i(t) (j = 1, 2, \dots, n)$ and a residue EMD, $r^i(t)$.

$$x^i(t) = \sum_{j=1}^n c_j^i(t) + r^i(t), \quad (2)$$

3. Repeat steps 1 and 2 M-times with different white noises until you get the appropriate decomposition result.
4. Calculate the average of the IMF trials which conducted by M-times as the final IMF.

$$c_j(t) = \frac{1}{M} \sum_{j=1}^n c_j^i(t), \quad (3)$$

The result of EEMD can be described as a linear combination of IMF and residuals formulated by Equation (4).

$$x(t) = \sum_{j=1}^n c_j(t) + r(t) \quad (4)$$

The amount of IMF denoted by n , $c_j(t)$ ($j = 1, 2, \dots, t$) is the j th IMF extracted in the j th decomposition in the interval t and $r(t)$ the final residue. It assumed that the complexity of the data involves information and noises, while the average of the ensemble data with different noises approximates the actual signal, white noises are used to capture the true IMF, and themselves are offset by the ensemble mean. The error of adding white noises can be controlled by Equation (5).

$$\varepsilon_{ne} = \frac{\varepsilon}{M}, \quad (5)$$

where:

M —number of ensemble members

ε —amplitude of the added noises series amplitude

ε_{ne} —final standard deviation of error

ε_{ne} —also interpreted as the difference between the input signal and the relevant IMF [32].

2.4. Wavelet Transformation

Wavelet transformation (WT) is a time-frequency decomposition method that provides a time series basis, both time and frequency [35]. Time series identical to non-stationary data. The WT is a wave that moves up and down in space-time periodically [36]. WT can also be described as a short wave. This model describes the frequency of the signal timing [37]. This model is also known as pre-processing for the denoising process on time series data by decomposing the data into several series [38].

WT identifies the signal shifts and then analyzes them. The purpose of the analysis is to obtain the information and frequency spectrum simultaneously [37]. This model is considered an effective model in analyzing time-frequency after special Fourier analysis in signal processing [39].

WT has two types of models, namely Continuous Wavelet Transformation (CWT) and Discrete Wavelet Transformation (DWT) [39]. CWT is modeled by Equation (6).

$$CWT_y(\alpha, \tau) = \frac{1}{\sqrt{|\alpha|}} \int y(t) \psi^* \left(\frac{t - \tau}{\alpha} \right) dt, \quad (6)$$

where α is parameters of scale, τ is translation parameters, $\psi^*(x)$ is conjugation complex functions, and $\psi(x)$ is mother wavelet.

The Discrete Wavelet Transformation (DWT) is modeled in Equation (7).

$$DWT_y(m, n) = \alpha_0^{-\frac{m}{2}} \int y(t) \psi^*(\alpha_0^{-m} t - n\tau_0) dt, \quad (7)$$

where m is a constant scale (decomposition level), and n is a constant integer.

3. Methodology

This study proposes a forecasting framework for electricity consumption. The main challenge in this problem is the complexity of the data pattern. Therefore, a decomposition algorithm is applied to understand data characteristics. Afterward, it applies an LSTM algorithm to forecast the electricity consumption. Figure 4 illustrates the proposed forecasting framework. The main contribution of this paper is the forecasting framework, which consists of decomposition algorithm. The decomposition algorithm divides time-series into several series. Thus, it needs an additional procedure to make a prediction for the data. Each series is predicted using neural network-based algorithm. Afterward, the results need an aggregation process. The proposed forecasting framework also involves the aggregation step for the data.

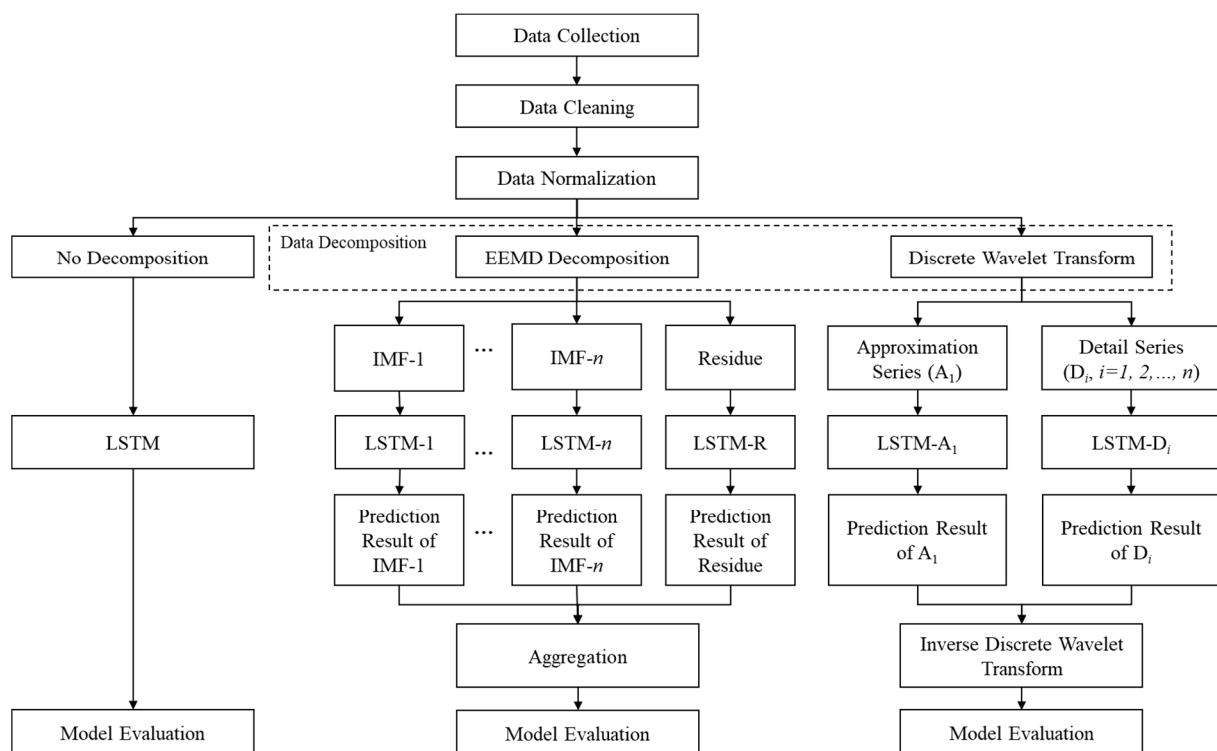


Figure 4. The proposed forecasting framework.

3.1. Data Pre-Processing

Data pre-processing is a critical part of processing complex time-series data. Therefore, the proposed framework applies data pre-processing techniques, including normalization, statistics descriptive analysis, and decomposition.

3.1.1. Data Normalization

Data normalization aims to avoid a large variety of data values. Normalization is carried out using the standard scaler method, as shown in Equation (8).

$$x_{scaled} = \frac{x_i - x_{mean}}{standard\ deviation} \quad (8)$$

3.1.2. Statistics Descriptive Analysis

Data exploration includes analyzing the data distribution by observing the data patterns, performing descriptive statistical tests, and performing stationarity tests using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and augmented Dickey-Fuller (ADF) to see the stationarity of the data to strengthen the theory of the method used. Furthermore, data

visualization is also applied to analyze the data before forecasting based on the pattern captured from the data visualization.

3.1.3. Data Decomposition

Decomposition is useful for data that have a high level of complexity and non-stationary data to improve the performance of the forecasting method that will be used. In the decomposition, two methods are used. They are WT and EEMD.

1. Wavelet Transformation

This study uses DWT to analyze the data with a time-scale function. WT has a multi-scale resolution and has time-shifting characteristics. Moreover, scaling operations can observe signals at different scales [40]. Thus, this method is very suitable for handling non-stationary time series data.

Based on Figure 4, the data are divided into two parts, training set $x_{train}(t) = 1, 2, \dots, n_{train}$ and testing set $x_{test}(t) = 1, 2, \dots, n_{test}$. The training set is used to build the forecasting model after the WT step. The WT has two components, namely approximation, and detailed series. The approximation series captures the low-frequency features, while the detailed series captures the high-frequency features of the original data. The approximation series is a further decomposed process with WT. The high-frequency noise representing fluctuations and irregularities is extracted and filtered [41]. DWT consists of two basic components of wavelet functions, namely father wavelet φ and mother wavelet ψ as shown in Equations (9) and (10) [38].

$$\varphi_{j,k}(x) = 2^{-\frac{j}{2}} \varphi(2^j t - k), \quad (9)$$

$$\psi_{j,k}(x) = 2^{-\frac{j}{2}} \psi(2^j t - k), \quad (10)$$

where $j = 1, \dots, J$ is the scaling parameter on the decomposition of each j th level. The k is the translation parameter. The approximation series is transformed by the father wavelet from the original data. The detailed series is transformed by the mother wavelet. The detailed series deals with oscillations of length $2^j - 2^{j+1}$. The approximation and detail series are modeled in Equations (11) and (12).

$$D_{j,t} = \int_{-\infty}^{\infty} y(x) \varphi_{j,t}(x) dx, \quad (11)$$

$$A_{j,t} = \int_{-\infty}^{\infty} y(x) \psi_{j,t}(x) dx \quad (12)$$

Finally, the time-series data on electricity consumptions consists of approximation and detail series as represented in Equation (13).

$$y(x) = A_j(x) + D_j(x) + D_{j-1}(x) + \dots + D_1(x), \quad (13)$$

2. Ensemble Empirical Mode Decomposition

The decomposition using EEMD is conducted as follows [30,42]:

- (1) Add the White Gaussian Noise series $\varepsilon_j(t)$ into the train set $x_{train}(t)$ and become the new series of $x_{train\ j}(t)$.
- (2) Decomposed the $x_{train\ j}(t)$ into several IMFs $c_j(t)$, $j = 1, 2, \dots, n_{train}$ and a residue $r(t)$.
- (3) Then, repeated the steps 1 and 2 on each $j = 1, 2, \dots, N_E$ by adding the white Gaussian noise series for every repeated process. N_E are the amount of repetition.
- (4) Take the average of all IMFs and the average of the residue as the final results.
- (5) The time series data after EEMD is the sum of all IMFs components and residues as shown in Equation (14).

$$x_{train}(t) = \sum_{j=1}^n c_j(t) + r(t), \quad (14)$$

3.2. Forecasting Using LSTM Algorithm

The proposed forecasting framework applies many-to-one RNN structure where each node in the hidden node uses the LSTM structure. Figure 5 illustrates the framework of RNN and the LSTM features.

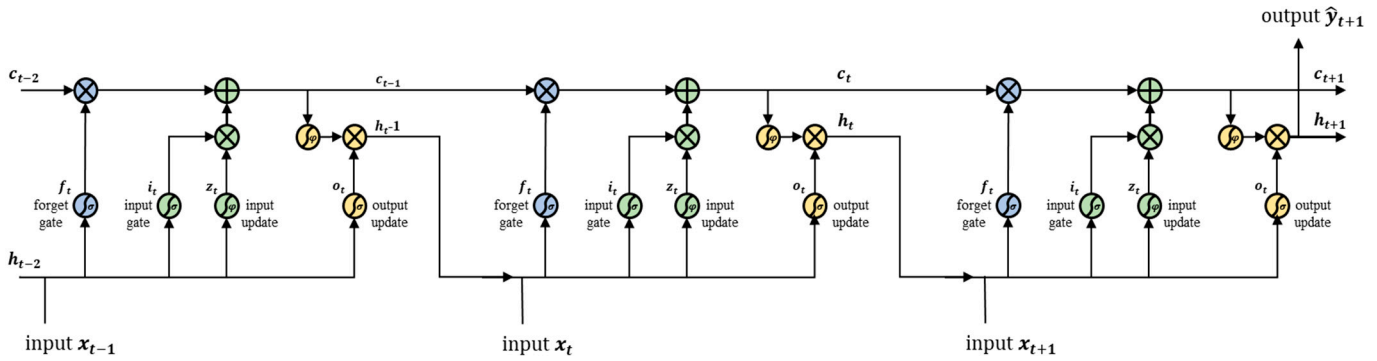


Figure 5. The RNN and LSTM structures.

The LSTM model consists of several stages [43]. It starts by determining the information that can pass in the cell state. The decision is controlled by the forget gate for time t , f_t , as shown in Equation (15).

$$f_t = \sigma(W^{fx}x_t + W^{fh}h_{t-1} + b_f), \quad (15)$$

where x_t is the input of time t , h_{t-1} is the output from the hidden node at time $t - 1$, W^{fx} is the weight between nodes in the input layer and forget gate, W^{fh} is the weight between nodes in the hidden layer and forget gate, and b_f is bias in forget gate.

The next step is determining information that must be entered into the cell state using Equation (16).

$$i_t = \sigma(W^{ix}x_t + W^{ih}h_{t-1} + b_i), \quad (16)$$

where the i_t is the output from input gate at time t , W^{ix} is the weight between nodes in the input layer and input gate, W^{ih} is the weight between nodes in the hidden layer and input gate, and b_i is bias in the input gate.

In addition, the candidate $\tilde{C}_t$ is generated using Equation (17).

$$\tilde{C}_t = \tanh(W^{cx}x_t + W^{ch}h_{t-1} + b_c), \quad (17)$$

where W^{cx} is the weight between nodes in the input layer to the output cell, W^{ch} is the weight between nodes in the hidden layer to the output cell, and b_c is the bias of the cell.

Forget the unwanted information by multiplying the old cell state C_{t-1} with f_t and adding some new information to the cell state by $i_t\tilde{C}_t$ as shown in Equation (18).

$$C_t = f_tC_{t-1} + i_t\tilde{C}_t, \quad (18)$$

Calculate the final output by compressing the tanh layer with C_t then multiplying by the output gate o_t as shown by Equations (19) and (20).

$$o_t = \sigma(W^{ox}x_t + W^{oh}h_{t-1} + b_o), \quad (19)$$

$$h_t = o_t \tanh(C_t) \quad (20)$$

4. Experimental Result

4.1. Dataset Description

This study applies the proposed forecasting framework to six benchmark datasets and one dataset taken from an education building in Taiwan. The six benchmark datasets consist of 16 buildings from different locations with different time zones and building functionalities. The dataset from Taiwan consists of four buildings data. There are 20 buildings studied in this research. These varieties make the data have different patterns and characteristics. The benchmark dataset consists of industrial buildings, education, commercial building, government, and residential. These data are available at <https://github.com/buds-lab/the-building-data-genome-project> [44] (accessed on 13 March 2021) and <http://traces.cs.umass.edu/index.php/Smart/Smart> [45] (accessed on 13 March 2021). The data were initially collected by using a building electricity meter [44]. First, the data are used to analyze the building in question. Then the data are extracted every certain period, one of which is the analysis of the energy system of the building [44]. The unit is a kilowatt-hour (kWh), and each dataset is distinguished based on the level of consumption. Table 2 shows the data specifications. The data are available as hourly data. The forecasting aims to predict one hour. In managing electricity consumption in a building, time hour interval is enough to set the devices.

Table 2. Dataset specifications.

Dataset	Units	Data Start	Data End	Functionalities	Timezone
1	UnivClass_Boyd	2012-01-01 0:00	2012-12-13 23:00	College Classroom	America/Los_Angeles
	UnivLab_Bethany	2012-01-01 0:00	2012-12-13 23:00	College Laboratory	America/Los_Angeles
	Office_Evelyn	2012-01-01 0:00	2012-12-13 23:00	Office	America/Los_Angeles
	Office_Bobbi	2012-01-01 0:00	2012-12-13 23:00	Office	America/Los_Angeles
2	UnivDorm_Malachi	2014-05-01 0:00	2015-04-30 23:00	Residential	America/Chicago
	UnivDorm_Mitch	2014-05-01 0:00	2015-04-30 23:00	Residential	America/Chicago
3	UnivClass_Seb	2014-12-01 0:00	2015-11-30 23:00	College Classroom	Europe/London
	UnivLab_Susan	2014-12-01 0:00	2015-11-30 23:00	College Laboratory	Europe/London
	Office_Stella	2014-12-01 0:00	2015-11-30 23:00	Office	Europe/London
	Office_Glenn	2014-12-01 0:00	2015-11-30 23:00	Office	Europe/London
4	Apt_Moon	2015-01-01 0:00	2015-12-31 23:00	Apartment	America/New_York
	Apt_Phobos	2015-01-01 0:00	2015-12-31 23:00	Apartment	America/New_York
5	UnivExhibit_Hermes	2019-01-02 0:00	2020-01-31 23:00	Exhibition	Asia/Taiwan
	UnivDorm_Athena	2019-01-02 0:00	2020-01-31 23:00	Residential	Asia/ Taiwan
	UnivMulti_Zeus	2019-01-02 0:00	2020-01-31 23:00	Multipurpose	Asia/ Taiwan
	UnivMulti_Ares	2019-01-02 0:00	2020-01-31 23:00	Multipurpose	Asia/ Taiwan
6	PrimClass_Uma	2012-02-02 0:00	2013-01-31 23:00	Primary/Secondary Classroom	Asia/Singapore
	PrimClass_Umar	2012-02-02 0:00	2013-01-31 23:00	Primary/Secondary Classroom	Asia/Singapore
	UnivDorm_Una	2012-02-02 0:00	2013-01-31 23:00	Residential	Asia/Singapore
	PrimClass_Ulysses	2012-02-02 0:00	2013-01-31 23:00	Primary/Secondary Classroom	Asia/Singapore

Source: [44,45], and an education building in Taiwan.

4.2. Statistic Description Analysis

Data exploration using a statistics description approach is a simple yet effective way to analyze the data concentration and distribution. Table 3 summarizes the statistics descriptive of the data. Table 3 represents the level of energy consumption of each building. The average value shows that dataset 5, the education industry, has very high consumption because it is used for residential in the University, multipurpose buildings, laboratories, classrooms, and offices. Multipurpose and residential buildings in the University have the highest average energy consumption. On the contrary, the apartment and primary/secondary class buildings have the lowest energy consumption.

Table 3. Descriptive statistics.

Dataset	Units	Mean	Stdev	Variance	Min	Max
1	UnivClass_Boyd	22.199	1.912	3.655	16.788	28.452
	UnivLab_Bethany	72.946	17.247	297.451	43.634	124.788
	Office_Evelyn	217.398	85.811	7,363.45	98.444	479.186
	Office_Bobbi	74.42	28.538	814.43	33.225	157.675
2	UnivDorm_Malachi	88.061	31.071	965.429	35.25	180.25
	UnivDorm_Mitch	62.82	17.281	298.637	32.25	116.25
3	UnivClass_Seb	63.365	25.922	671.973	10	141
	UnivLab_Susan	19.157	11.801	139.27	4	56
	Office_Stella	66.442	26.473	700.829	20	143
	Office_Glenn	34.035	15.484	239.753	5.1	85.26
4	Apt_Moon	1.077	0.93	0.864	0.018	6.872
	Apt_Phobos	1.382	1.241	1.539	0.012	8.715
5	UnivExhibit_Hermes	0.293	0.240	0.058	0.044	2.68
	UnivDorm_Athena	111.908	46.138	2128.691	26.246	347.304
	UnivMulti_Zeus	90.112	33.631	1131.061	35.927	222.969
	UnivMulti_Ares	28.454	13.916	193.655	9.055	94.036
6	PrimClass_Uma	12.375	9.01	81.184	1	39
	PrimClass_Umar	21.739	17.24	297.221	1	73
	UnivDorm_Una	23.535	8.673	75.213	1	49
	PrimClass_Ulysses	25.279	26.157	684.207	1	103

Furthermore, a data stationarity test is carried out to show whether each series is stationary or non-stationary. This study performs the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test and the augmented Dickey-Fuller (ADF) test for the stationarity test. The KPSS test is a unit root test for the stationarity of data on a trend. It is often referred to as a trend-stationary test [46]. The ADF test is a statistical test that is more popular than the KPSS test [47]. The ADF test has the same function as the KPSS test, but it tests whether the data are linear or the difference is stationary. The difference between trend or difference stationary is that the stationary trend has a deterministic average, or the shocks caused to the data will return to its path. While difference stationary has a stochastic average property that has permanent shocks, it has the weakest stationary nature of the data [47].

The hypothesis for the KPSS test is as follows:

H_0 —The data are stationary (trend stationary)

H_1 —The data have a unit root. Thus, the data are non-stationary

Table 4 shows the results of the KPSS test, which consists of the p-value and the critical value using four different settings. If the p-value < 0.05, then we reject H_0 , which means

the data are non-stationary. Otherwise, the data are stationary. Based on the KPSS test, Office_Evelyn, UnivLab_Susan, and PrimClass_Umar are stationary.

Table 4. Results of KPSS and ADF test.

Series	KPSS Test			ADF Test			Combination of KPSS and ADF Results
	Test Statistics	<i>p</i> -Value	Interpretation	Test Statistics	<i>p</i> -Value	Interpretation	
UnivClass_Boyd	1.325	0.01	non-stationary	−7.268	1.61×10^{-10}	stationary	difference stationary
UnivLab_Bethany	0.319	0.01	non-stationary	−10.656	4.54×10^{-19}	stationary	difference stationary
Office_Evelyn	0.143	0.055	stationary	−13.55	2.42×10^{-25}	stationary	stationary
Office_Bobbi	0.301	0.01	non-stationary	−10.983	7.37×10^{-20}	stationary	difference stationary
UnivDorm_Malachi	1.19	0.01	non-stationary	−2.511	0.1128005	non-stationary	non-stationary
UnivDorm_Mitch	1.091	0.01	non-stationary	−3.878	0.0022008	stationary	difference stationary
UnivClass_Seb	0.715	0.01	non-stationary	−11.924	4.96×10^{-22}	stationary	difference stationary
UnivLab_Susan	0.087	0.1	stationary	−11.191	2.36×10^{-20}	stationary	stationary
Office_Stella	2.471	0.01	non-stationary	−5.587	1.36×10^{-6}	stationary	difference stationary
Office_Glenn	1.63	0.01	non-stationary	−5.76	5.70×10^{-7}	stationary	difference stationary
Apt_Moon	3.329	0.01	non-stationary	−4.209	0.000635	stationary	difference stationary
Apt_Phobos	3.636	0.01	non-stationary	−4.159	0.0007739	stationary	difference stationary
UnivExhibit_Hermes	1.966	0.01	non-stationary	−8.045	1.81×10^{-12}	stationary	difference stationary
UnivDorm_Athena	1.805	0.01	non-stationary	−4.727	7.49×10^{-5}	stationary	difference stationary
UnivMulti_Zeus	2.787	0.01	non-stationary	−6.364	2.43×10^{-8}	stationary	difference stationary
UnivMulti_Ares	1.663	0.01	non-stationary	−6.671	4.58×10^{-9}	stationary	difference stationary
PrimClass_Uma	0.232	0.01	non-stationary	−11.223	2.00×10^{-20}	stationary	difference stationary
PrimClass_Umar	0.119	0.1	stationary	−12.612	1.65×10^{-23}	stationary	stationary
UnivDorm_Una	0.577	0.01	non-stationary	−3.643	0.0049932	stationary	difference stationary
PrimClass_Ulysses	0.216	0.01	non-stationary	−11.645	2.11×10^{-21}	stationary	difference stationary

On the other hand, the hypothesis for the SDF test is as follows:

H_0 —the data have a unit root. Thus, the data are non-stationary

H_1 —the data are stationary (difference stationary)

Unlike the KPSS test, the hypothesis of the ADF test is the opposite. If the *p*-value < 0.05, then reject H_0 which means the data are stationary (difference stationary). Otherwise, the data are non-stationary. The ADF test result in Table 4 shows that only UnivDorm_Malachi is non-stationary.

To see the level of the stationarity of data based on the KPSS and ADF tests, this study refers to the possible outputs generated based on:

- Case 1: KPSS and ADF tests are non-stationary, then the series is non-stationary.
- Case 2: KPSS and ADF tests are stationary, so the series is stationary.

- Case 3: KPSS test is stationary, and ADF is non-stationary, then the series is trend-stationary. The treatment must be done by eliminating the trend to make the series stationary.
- Case 4: KPSS is non-stationary, and ADF is stationary, so the series the difference stationary. Differentiation is used to create a stationary series.

The combined result of the KPSS and ADF tests in Table 4 reveals that UnivDorm_Malachi is non-stationary. The Office_Evelyn, UnivLab_Susan, and PrimeClass_Umar are stationary. The other series is the difference stationary category (weak), which is the lowest stationary category.

Data visualization in Figure 6 presents the pattern of each series. The charts are available in a span of one year of hourly data resolution. Figure 6 shows that several buildings have similarities in data patterns. Some buildings have similar consumption patterns as Apt_Moon and Apt_Phobos buildings, but both buildings have very low consumption levels. In addition, UnivDorm_Malachi, as the only non-stationary series, has an uneven consumption pattern. Some buildings have stable consumption patterns such as Office_Bobbi, Office_Evelyn, UnivClass_Boyd, UnivLab_Bethany, UnivDorm_Mitch, Office_Glenn, Office_Stella, UnivClass_Seb, UnivLab_Susan, UnivMulti_Ares, UnivMulti_Zeus, PrimClass_Ulysses, PrimClass_Ulysses. Buildings UnivDorm_Athena, UnivExhibit_Hermes, UnivMulti_Ares, UnivMulti_Zeus have a very high consumption rate and the highest among other buildings.

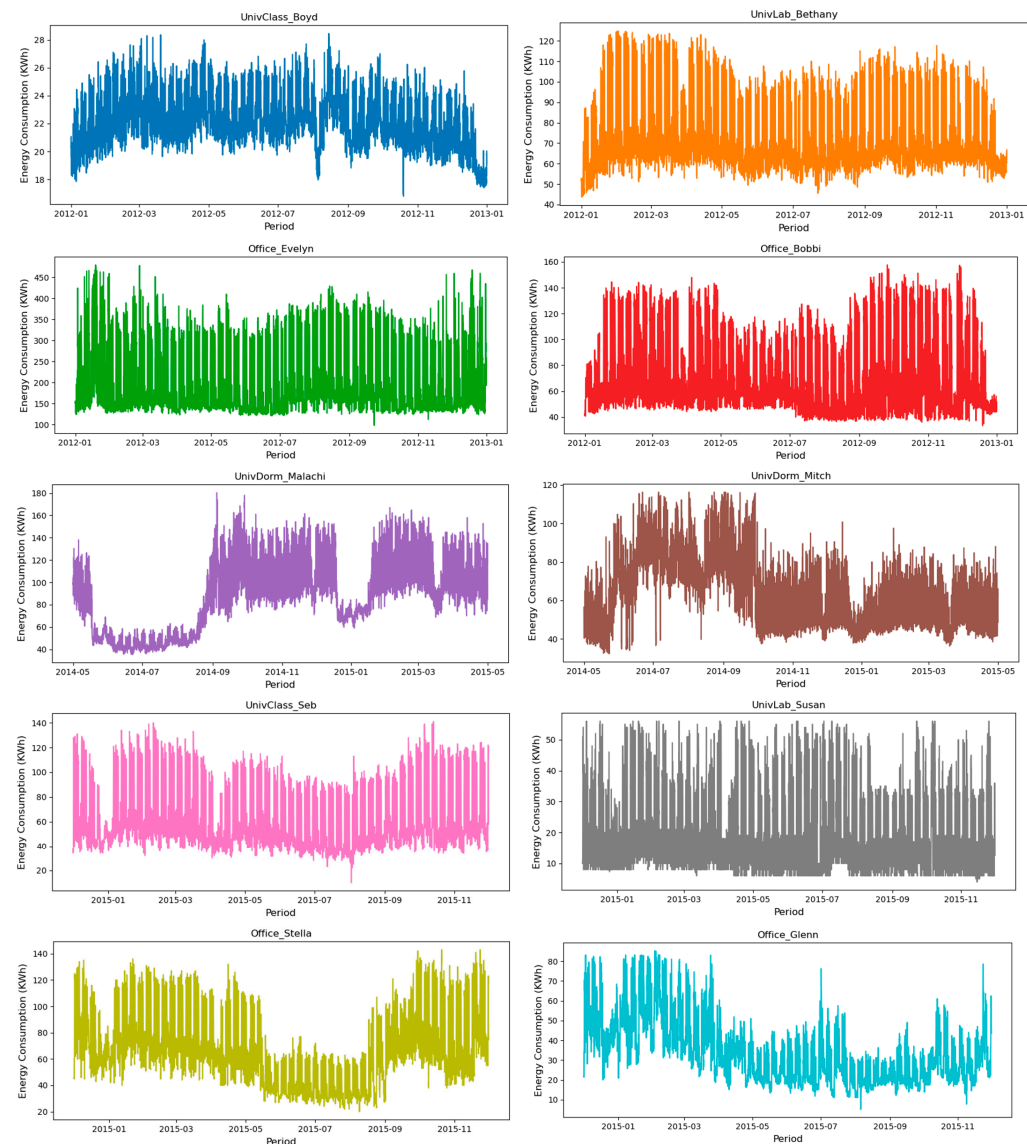


Figure 6. Cont.

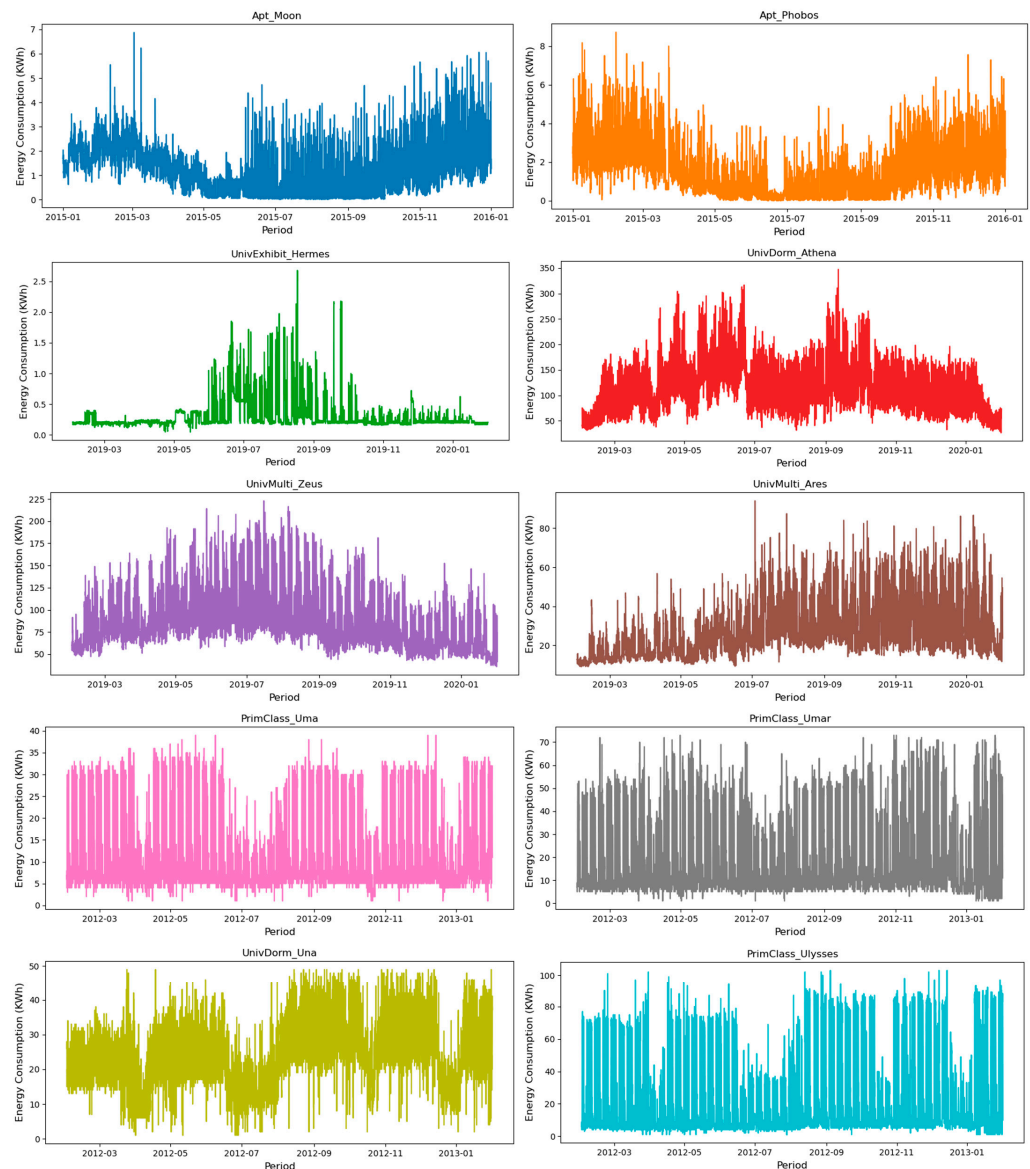


Figure 6. Energy consumption pattern in each building.

4.3. Parameter Setting

This study applies the proposed forecasting framework to predict one hour ahead data based on one-year data for each building. Herein, the data period is hourly. In the forecasting process, the data are divided into training, validation, and testing data. Training data are 70% of the data population of each building, validation data are 20%, and testing data are 10%. The data splitting process is carried out sequentially, accounting for 70% of the first data for training. Then, after 70% of the training data, 20% of the data is used for validation, and the remaining 10% of the data for testing data is calculated after 20% of the validation data until the last data. The validation is the result of the splitting of data testing carried out to overcome overfitting in the data, and the model does not understand the character of the data [48].

The LSTM involves several parameters, including the number of hidden nodes, look-back periods, activation function, and recurrent activation. To get the best parameter setting for the forecasting model, this study evaluates several parameter settings as listed in Table 5. Each parameter set is used for all buildings, with five replications for each building.

Table 5. Trial of parameter combinations.

Combinations	Hidden Nodes	Look Back	Activation	Recurrent Activation
1	64	24	sigmoid	sigmoid
2	64	24	sigmoid	relu
3	64	24	sigmoid	tanh
4	64	24	tanh	Sigmoid
5	64	24	tanh	relu
6	64	24	tanh	tanh
7	64	24	relu	sigmoid
8	64	24	relu	tanh
9	64	168	sigmoid	sigmoid
10	64	168	sigmoid	tanh
11	64	168	tanh	Sigmoid
12	64	168	tanh	tanh
13	128	24	sigmoid	sigmoid
14	128	24	sigmoid	relu
15	128	24	sigmoid	tanh
16	128	24	tanh	Sigmoid
17	128	24	tanh	relu
18	128	24	tanh	tanh
19	128	24	relu	sigmoid
20	128	24	relu	tanh
21	128	168	sigmoid	sigmoid
22	128	168	sigmoid	tanh
23	128	168	tanh	Sigmoid
24	128	168	tanh	tanh

Table 6 shows the best parameter setting for LSTM. This parameter is used of forecasting without decomposition, with EEMD-LSTM, and WT-LSTM. According to Table 6, the forecasting is conducted using 24 h of look back data. It means that predicting the next one hour requires the previous 24 h of historical data. The number of hidden nodes indicates the number of nodes contained in the hidden layers of the LSTM network. The optimizer is Nadam. Nadam is one of the adaptive movement estimations (Adam) that added momentum to the Nesterov accelerated gradient (NAG). This optimizer works for the gradient descent process in the process of finding the minimum function by increasing the speed of convergence and the quality of the studied model. Epoch is number of iterations used for the forward and backward process on developing the model based on the training dataset. Activation is used in the cell state of the LSTM structure, while recurrent activation is used for the gates in the LSTM including input gate, forget gate and output gate.

The forecasting model with EEMD-LSTM involves more parameters. They are the noise width (white noise's standard deviation) and ensemble member (trial). This study uses 0.05 as the noise width and 100 for the ensemble member [32]. Furthermore, the WT-LSTM is set using a single level of decomposition, symmetric signal extension model, and Daubechies (db1) wavelet object.

Table 6. Best parameters setting for the LSTM.

Parameters	Values
Layers	1 layer
Hidden nodes	64 nodes
Look back	24 h
Epochs	50 epochs
Optimizer	Nadam
Activation	Tanh
Recurrent activation	Sigmoid

4.4. Forecasting Results Analysis

To evaluate the robustness of the proposed forecasting framework, each forecasting method is run ten times for each data series. This study evaluates the results based on the mean absolute percentage error (*MAPE*), mean squared error (*MSE*), and mean absolute error (*MAE*). These parameters are calculated using Equations (21)–(23).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (21)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (22)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (23)$$

where y_i is the actual value of the period i , $\hat{y}_i$ is the predicted value of the period i , and n is the number of periods. To evaluate the performance of the LSTM, this study also compares the LSTM with the recurrent neural network (RNN) algorithm.

Tables 7–10 shows the average and standard deviation value of the results. Table 11 shows the comparison of the three approaches based on the *MAPE* value. These results reveal that the eight different buildings, including office buildings, lecture laboratories, lecture classrooms, dormitory or residential in university areas, apartments, multipurpose, and secondary classrooms, have very error rates. The training error of the EEMD-LSTM method has the lowest average *MAPE* value compared to the LSTM and WT-LSTM. It is followed by the testing error of the EEMD-LSTM, which is still the lowest among the other methods. It can be caused by the decomposition method itself, which helps in the denoising process of the time-series data used.

Figure 7 shows the comparison between the actual and predicted value of the energy consumption. It reveals that for most of the buildings, the predicted value is very close to the actual value. Furthermore, the standard deviation of the EEMD-LSTM algorithm shown in Table 9 is relatively small. It indicates that the EEMD-LSTM algorithm has a stable result.

Table 7. The MAPE, MSE, and MAE of forecasting using LSTM without decomposition.

Dataset	MAPE (%)				MAE			MSE		
	Metrics	Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	1.95	1.89	2.08	0.44	0.43	0.43	0.34	0.31	0.32
	std	0.02	0.02	0.10	0.00	0.01	0.02	0.00	0.01	0.02
Univlab_Bethany	mean	2.78	3.00	2.80	2.11	2.19	2.05	9.14	9.67	8.58
	std	0.04	0.04	0.04	0.03	0.03	0.03	0.20	0.42	0.31
Office_Evelyn	mean	4.61	4.63	6.01	10.05	10.10	13.12	438.18	352.39	674.85
	std	0.14	0.32	0.12	0.28	0.71	0.23	17.13	48.76	17.05
Office_Bobbi	mean	5.72	7.87	7.83	4.29	5.18	5.40	36.75	48.39	53.76
	std	0.08	0.25	0.17	0.05	0.13	0.11	0.90	2.03	2.06
Univdorm_Malachi	mean	4.76	4.71	5.77	3.78	4.86	6.32	35.43	47.07	68.26
	std	0.08	0.11	0.12	0.05	0.08	0.10	0.55	0.85	1.99
Univdorm_Mitch	mean	8.10	7.98	7.85	5.74	4.55	4.53	63.85	39.92	38.46
	std	0.08	0.31	0.19	0.04	0.18	0.12	0.85	2.65	2.42
Univclass_Seb	mean	3.64	4.75	3.97	2.18	2.40	2.58	8.64	11.63	12.99
	std	0.09	0.25	0.11	0.05	0.12	0.07	0.39	1.02	0.86
Univlab_Susan	mean	7.60	13.65	11.11	1.68	2.60	1.97	9.89	21.04	11.58
	std	0.58	0.67	0.85	0.07	0.09	0.12	0.46	0.92	0.71
Office_Stella	mean	3.49	5.32	4.07	2.36	2.26	3.12	12.39	11.29	19.02
	std	0.07	0.27	0.12	0.05	0.07	0.10	0.63	0.44	1.11
Office_Glenn	mean	6.77	6.71	7.71	2.84	1.63	2.09	16.74	5.80	8.60
	std	0.10	0.36	0.16	0.03	0.07	0.03	0.18	0.31	0.20
Apt_Moon	mean	34.31	128.03	39.38	0.21	0.34	0.54	0.16	0.43	0.73
	std	2.32	10.51	0.93	0.00	0.00	0.01	0.00	0.01	0.01
Apt_Phobos	mean	47.92	116.12	49.06	0.42	0.38	0.76	0.58	0.40	1.09
	std	5.82	13.42	1.26	0.00	0.00	0.00	0.01	0.01	0.01
Univexhibit_Hermes	mean	5.95	9.27	4.98	0.02	0.05	0.01	0.01	0.02	0.00
	std	0.12	0.27	0.16	0.00	0.00	0.00	0.00	0.00	0.00
Univdorm_Athena	mean	8.05	8.35	9.69	8.40	9.33	8.32	120.67	155.49	117.45
	std	0.12	0.19	0.24	0.12	0.24	0.13	3.35	7.31	3.10
Univmulti_Zeus	mean	5.29	5.66	8.02	5.15	5.18	5.42	53.82	57.52	66.14
	std	0.10	0.20	0.68	0.05	0.13	0.32	0.54	2.40	2.78
Univmulti_Ares	mean	7.36	11.21	15.42	1.74	4.01	5.38	9.30	34.07	57.41
	std	0.17	0.19	0.25	0.03	0.05	0.07	0.20	1.02	1.33
Primclass_Uma	mean	19.69	17.90	17.27	1.53	1.58	1.56	10.11	8.45	9.41
	std	0.45	0.64	0.56	0.03	0.04	0.04	0.28	0.41	0.46
Primclass_Umar	mean	20.59	14.96	24.42	2.67	2.73	3.84	29.92	24.17	41.45
	std	0.73	0.60	0.71	0.07	0.07	0.07	0.79	1.40	1.16
Univdorm_Una	mean	18.33	15.33	16.71	2.29	3.64	3.35	15.76	28.33	25.60
	std	0.43	0.69	0.36	0.04	0.19	0.10	0.49	3.15	1.63
Primclass_Ulysses	mean	27.09	22.96	28.85	3.33	4.67	5.24	66.45	88.19	99.04
	std	0.91	0.77	0.65	0.09	0.19	0.12	1.44	4.82	2.22

Table 8. The MAPE, MSE, and MAE of forecasting using RNN without decomposition.

Dataset	MAPE (%)				MAE			MSE		
	Metrics	Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	1.76	2.05	2.20	0.40	0.47	0.46	0.29	0.36	0.35
	std	0.03	0.02	0.05	0.01	0.01	0.01	0.01	0.00	0.01
Univlab_Bethany	mean	2.73	3.03	2.94	2.06	2.21	2.15	8.77	9.51	8.99
	std	0.08	0.08	0.07	0.05	0.05	0.05	0.31	0.41	0.53
Office_Evelyn	mean	4.54	4.07	6.40	9.69	8.70	13.80	401.08	192.37	692.24
	std	0.10	0.25	0.14	0.24	0.64	0.25	13.18	24.36	27.48
Office_Bobbi	mean	5.47	8.10	8.02	4.07	5.27	5.50	33.29	48.61	54.71
	std	0.13	0.47	0.32	0.07	0.20	0.15	0.90	2.75	2.25
Univdorm_Malachi	mean	4.75	4.94	6.05	3.71	5.12	6.65	33.93	51.35	75.23
	std	0.16	0.15	0.14	0.08	0.13	0.14	1.05	1.79	2.95
Univdorm_Mitch	mean	7.51	8.51	8.39	5.27	4.83	4.81	56.00	44.73	42.79
	std	0.14	0.28	0.11	0.09	0.15	0.07	1.52	2.30	1.23
Univclass_Seb	mean	3.69	4.93	4.11	2.23	2.51	2.68	8.97	12.88	14.04
	std	0.27	0.26	0.26	0.15	0.12	0.15	0.94	0.98	1.09
Univlab_Susan	mean	7.76	13.98	10.54	1.66	2.59	1.88	9.15	20.25	10.53
	std	0.35	0.81	0.66	0.04	0.07	0.06	0.35	0.64	0.46
Office_Stella	mean	3.33	5.37	4.02	2.24	2.25	3.09	10.67	10.93	18.45
	std	0.08	0.32	0.10	0.04	0.11	0.08	0.38	0.77	0.72
Office_Glenn	mean	6.54	7.52	8.34	2.69	1.77	2.23	15.33	6.48	9.48
	std	0.33	1.09	0.58	0.10	0.18	0.11	0.75	0.68	0.50
Apt_Moon	mean	36.55	159.74	42.12	0.21	0.36	0.55	0.15	0.43	0.74
	std	2.59	16.61	0.85	0.01	0.01	0.01	0.01	0.01	0.01
Apt_Phobos	mean	54.35	151.33	52.99	0.41	0.39	0.80	0.53	0.39	1.13
	std	8.97	23.68	1.77	0.01	0.01	0.01	0.02	0.02	0.03
Univexhibit_Hermes	mean	6.15	9.73	5.44	0.03	0.05	0.01	0.01	0.02	0.00
	std	0.11	0.15	0.17	0.00	0.00	0.00	0.00	0.00	0.00
Univdorm_Athena	mean	7.63	8.48	10.13	7.87	9.46	8.73	108.82	159.33	127.89
	std	0.34	0.29	0.47	0.22	0.31	0.26	4.32	9.61	7.01
Univmulti_Zeus	mean	5.21	5.79	7.84	5.04	5.30	5.21	52.10	58.06	63.32
	std	0.13	0.22	0.57	0.12	0.16	0.27	1.91	2.34	2.67
Univmulti_Ares	mean	7.35	11.59	15.58	1.70	4.12	5.44	8.73	34.78	57.09
	std	0.41	0.23	0.23	0.07	0.07	0.08	0.29	0.91	1.47
Primclass_Uma	mean	18.71	18.13	17.58	1.49	1.66	1.62	9.68	9.39	9.71
	std	0.66	0.86	0.60	0.05	0.05	0.04	0.38	0.40	0.39
Primclass_Umar	mean	19.80	15.31	26.64	2.56	2.78	3.94	28.12	24.25	42.46
	std	1.02	0.73	1.23	0.10	0.11	0.13	0.56	1.09	1.24
Univdorm_Una	mean	16.60	15.87	18.10	2.00	3.78	3.63	12.95	28.49	28.56
	std	0.55	0.64	0.45	0.05	0.18	0.11	0.57	2.29	1.50
Primclass_Ulysses	mean	26.55	23.11	31.51	3.22	4.71	5.32	62.17	88.43	98.03
	std	1.62	0.90	1.17	0.12	0.20	0.18	1.93	3.71	3.50

Table 9. The MAPE, MSE, and MAE of forecasting using EEMD-LSTM model.

Dataset	MAPE (%)				MAE			MSE		
	Metrics	Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	0.63	0.71	1.66	0.14	0.16	0.34	0.04	0.04	0.19
	std	0.01	0.04	0.12	0.00	0.01	0.02	0.00	0.00	0.03
Univlab_Bethany	mean	1.00	1.09	1.05	0.74	0.78	0.76	1.02	1.06	1.00
	std	0.02	0.01	0.01	0.01	0.00	0.00	0.04	0.01	0.01
Office_Evelyn	mean	1.99	1.47	2.65	4.01	3.08	5.56	49.55	17.35	113.79
	std	0.04	0.01	0.02	0.08	0.03	0.02	3.20	0.24	1.50
Office_Bobbi	mean	1.90	2.30	2.43	1.36	1.50	1.65	3.25	3.85	4.90
	std	0.04	0.02	0.01	0.03	0.02	0.01	0.14	0.07	0.06
Univdorm_Malachi	mean	1.95	2.36	4.80	1.43	2.22	4.97	4.22	7.71	31.57
	std	0.17	0.45	1.15	0.11	0.35	1.18	0.51	1.83	12.41
Univdorm_Mitch	mean	2.75	3.00	3.52	1.89	1.64	1.95	7.58	4.96	6.12
	std	0.04	0.18	0.22	0.04	0.09	0.12	0.29	0.47	0.50
Univclass_Seb	mean	1.41	1.72	1.56	0.83	0.83	0.98	1.16	1.17	1.98
	std	0.02	0.02	0.02	0.01	0.01	0.01	0.04	0.02	0.07
Univlab_Susan	mean	4.40	7.02	9.23	0.84	1.23	1.29	1.94	4.70	2.79
	std	0.26	0.37	0.95	0.06	0.05	0.12	0.26	0.07	0.34
Office_Stella	mean	1.33	2.04	1.50	0.91	0.82	1.14	1.40	1.18	2.20
	std	0.07	0.07	0.10	0.05	0.02	0.08	0.18	0.05	0.28
Office_Glenn	mean	2.38	2.98	3.24	0.94	0.68	0.85	1.69	0.75	1.12
	std	0.07	0.15	0.53	0.03	0.03	0.13	0.19	0.07	0.30
Apt_Moon	mean	23.93	118.99	21.52	0.08	0.16	0.24	0.02	0.07	0.13
	std	1.26	0.86	0.36	0.00	0.00	0.00	0.00	0.00	0.00
Apt_Phobos	mean	34.12	70.89	19.59	0.17	0.14	0.30	0.07	0.04	0.18
	std	1.14	1.72	0.11	0.00	0.00	0.01	0.00	0.00	0.01
Univexhibit_Hermes	mean	5.24	7.91	5.20	0.01	0.02	0.01	0.00	0.00	0.00
	std	0.15	0.12	0.23	0.00	0.00	0.00	0.00	0.00	0.00
Univdorm_Athena	mean	2.79	2.58	5.25	2.85	2.88	4.40	13.05	13.86	28.51
	std	0.05	0.02	1.15	0.06	0.04	0.86	0.47	0.44	10.02
Univmulti_Zeus	mean	1.83	2.53	3.11	1.71	2.20	1.98	5.26	7.98	8.44
	std	0.01	0.16	0.22	0.02	0.12	0.16	0.15	0.70	0.54
Univmulti_Ares	mean	3.39	4.52	6.22	0.73	1.55	1.98	1.19	5.00	7.10
	std	0.08	0.10	0.33	0.02	0.02	0.07	0.07	0.08	0.38
Primclass_Uma	mean	7.03	6.76	6.84	0.65	0.67	0.69	1.67	1.49	1.48
	std	0.10	0.09	0.11	0.00	0.01	0.02	0.02	0.06	0.09
Primclass_Umar	mean	8.30	6.84	9.60	1.08	1.09	1.52	3.73	3.72	6.72
	std	0.21	0.14	0.18	0.03	0.01	0.01	0.23	0.08	0.10
Univdorm_Una	mean	5.30	4.99	5.52	0.88	1.32	1.28	2.42	4.18	4.51
	std	0.12	0.06	0.25	0.02	0.01	0.04	0.09	0.25	0.22
Primclass_Ulysses	mean	11.97	10.66	15.31	1.41	1.75	2.08	7.87	10.13	13.39
	std	0.31	0.62	1.60	0.03	0.08	0.18	0.34	0.52	1.06

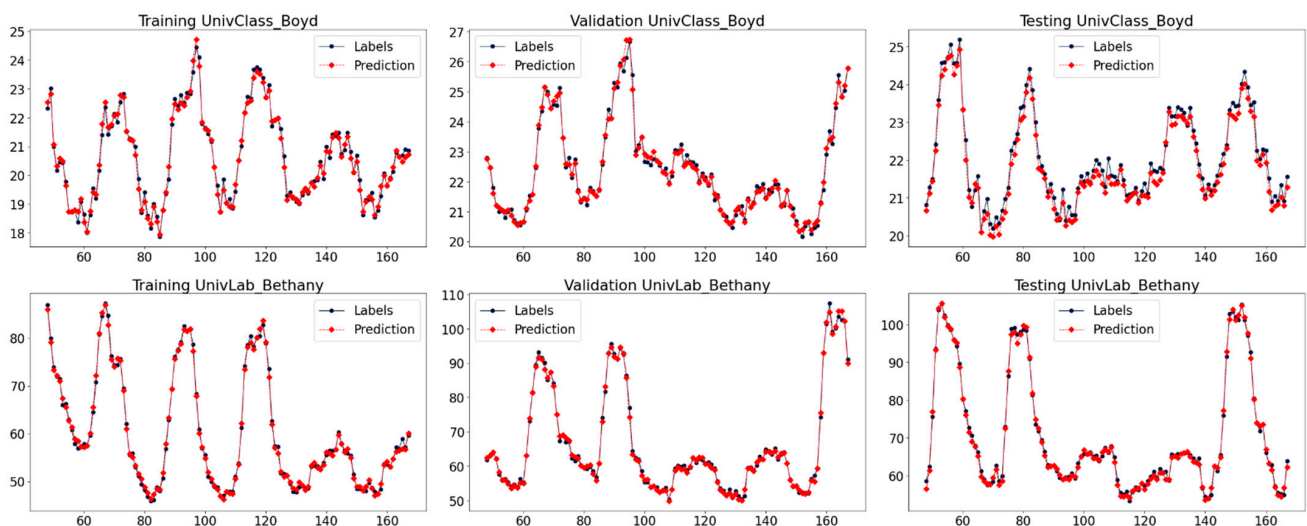
Table 10. The MAPE, MSE, and MAE of forecasting using WT-LSTM model.

Dataset	MAPE (%)				MAE			MSE		
	Metrics	Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	2.27	2.13	2.32	0.52	0.49	0.49	0.45	0.41	0.43
	std	0.03	0.03	0.15	0.01	0.01	0.03	0.01	0.01	0.04
Univlab_Bethany	mean	3.67	3.77	3.72	2.79	2.78	2.71	17.40	16.80	15.66
	std	0.13	0.06	0.08	0.09	0.04	0.02	1.40	0.64	0.61
Office_Evelyn	mean	6.93	6.46	8.52	14.94	13.87	18.26	778.93	577.21	1101.00
	std	0.28	0.14	0.35	0.12	0.18	0.26	27.29	48.76	16.85
Office_Bobbi	mean	7.13	8.96	9.24	5.31	5.87	6.36	53.01	64.22	76.37
	std	0.50	0.52	0.51	0.30	0.17	0.17	4.16	1.83	2.95
Univdorm_Malachi	mean	5.69	5.37	6.70	4.48	5.52	7.32	46.42	59.43	92.47
	std	0.12	0.04	0.11	0.11	0.02	0.16	1.74	0.43	4.79
Univdorm_Mitch	mean	8.95	8.18	8.12	6.22	4.63	4.64	71.12	40.07	39.44
	std	0.05	0.09	0.09	0.04	0.03	0.04	0.77	0.32	0.23
Univclass_Seb	mean	4.95	6.38	5.58	2.93	3.27	3.58	17.13	23.83	26.35
	std	0.29	0.07	0.15	0.16	0.03	0.12	1.90	0.56	2.36
Univlab_Susan	mean	9.11	16.59	14.42	2.04	3.22	2.57	13.30	29.66	17.03
	std	0.18	0.32	0.29	0.05	0.05	0.07	0.34	0.88	0.51
Office_Stella	mean	4.93	6.94	5.58	3.28	2.95	4.24	21.66	19.04	32.83
	std	0.27	0.20	0.15	0.19	0.06	0.12	1.56	0.28	1.34
Office_Glenn	mean	7.91	8.01	8.89	3.31	1.98	2.44	22.41	9.12	12.21
	std	0.01	0.15	0.05	0.01	0.01	0.01	0.10	0.07	0.05
Apt_Moon	mean	40.73	192.76	46.55	0.23	0.38	0.57	0.18	0.46	0.77
	std	0.44	3.32	0.48	0.00	0.00	0.00	0.00	0.00	0.00
Apt_Phobos	mean	67.95	184.62	59.32	0.48	0.40	0.82	0.68	0.40	1.17
	std	4.94	15.60	1.25	0.00	0.00	0.01	0.01	0.01	0.02
Univexhibit_Hermes	mean	10.06	15.13	7.55	0.04	0.07	0.02	0.01	0.03	0.00
	std	0.25	0.06	0.04	0.00	0.00	0.00	0.00	0.00	0.00
Univdorm_Athena	mean	9.68	10.20	12.47	9.92	11.25	10.68	163.75	229.93	189.17
	std	0.24	0.08	0.37	0.17	0.11	0.24	4.29	3.24	8.39
Univmulti_Zeus	mean	6.12	6.72	8.39	6.04	6.33	5.83	76.16	93.89	80.47
	std	0.07	0.05	0.04	0.06	0.03	0.05	0.92	0.90	1.85
Univmulti_Ares	mean	11.57	15.12	19.66	2.60	5.32	6.81	17.10	54.41	90.39
	std	0.27	0.14	0.50	0.05	0.02	0.13	0.43	0.58	3.07
Primclass_Uma	mean	19.61	18.80	18.36	1.69	1.77	1.81	11.61	12.21	12.83
	std	0.36	0.34	0.32	0.04	0.03	0.03	0.26	0.33	0.23
Primclass_Umar	mean	22.26	17.09	28.61	3.13	3.10	4.45	32.27	29.29	47.92
	std	0.81	0.43	0.49	0.08	0.03	0.01	1.14	0.72	0.16
Univdorm_Una	mean	21.54	17.96	18.51	2.92	4.31	3.82	19.63	33.75	28.21
	std	1.52	0.90	1.09	0.29	0.20	0.31	2.04	2.44	3.73
Primclass_Ulysses	mean	28.72	22.05	29.78	3.65	4.59	5.38	71.42	92.63	109.26
	std	0.98	0.50	0.38	0.18	0.11	0.10	1.13	2.14	2.52

Table 11. Comparison of algorithms based on MAPE (%).

Dataset	Training				Validation				Testing			
	RNN	LSTM	EEMD LSTM	WT LSTM	RNN	LSTM	EEMD LSTM	WT LSTM	RNN	LSTM	EEMD LSTM	WT LSTM
Univclass_Boyd	1.76	1.95	0.63 *	2.27	2.05	1.89	0.71 *	2.13	2.2	2.08	1.66 *	2.32
Univlab_Bethany	2.73	2.78	1 *	3.67	3.03	3	1.09 *	3.77	2.94	2.8	1.05 *	3.72
Office_Evelyn	4.54	4.61	1.99 *	6.93	4.07	4.63	1.47 *	6.46	6.4	6.01	2.65 *	8.52
Office_Bobbi	5.47	5.72	1.9 *	7.13	8.1	7.87	2.3 *	8.96	8.02	7.83	2.43 *	9.24
Univdorm_Malachi	4.75	4.76	1.95 *	5.69	4.94	4.71	2.36 *	5.37	6.05	5.77	4.8 *	6.7
Univdorm_Mitch	7.51	8.1	2.75 *	8.95	8.51	7.98	3 *	8.18	8.39	7.85	3.52 *	8.12
Univclass_Seb	3.69	3.64	1.41 *	4.95	4.93	4.75	1.72 *	6.38	4.11	3.97	1.56 *	5.58
Univlab_Susan	7.76	7.6	4.4 *	9.11	13.98	13.65	7.02 *	16.59	10.54	11.11	9.23 *	14.42
Office_Stella	3.33	3.49	1.33 *	4.93	5.37	5.32	2.04 *	6.94	4.02	4.07	1.5 *	5.58
Office_Glenn	6.54	6.77	2.38 *	7.91	7.52	6.71	2.98 *	8.01	8.34	7.71	3.24 *	8.89
Apt_Moon	36.55	34.31	23.93 *	40.73	159.74	128.03	118.9 *	192.76	42.12	39.38	21.52 *	46.55
Apt_Phobos	54.35	47.92	34.12 *	67.95	151.33	116.12	70.89 *	184.62	52.99	49.06	19.59 *	59.32
Univexhibit_Hermes	6.15	5.95	5.24 *	10.06	9.73	9.27	7.91 *	15.13	5.44	4.98 *	5.2	7.55
Univdorm_Athena	7.63	8.05	2.79 *	9.68	8.48	8.35	2.58 *	10.2	10.13	9.69	5.25 *	12.47
Univmulti_Zeus	5.21	5.29	1.83 *	6.12	5.79	5.66	2.53 *	6.72	7.84	8.02	3.11 *	8.39
Univmulti_Ares	7.35	7.36	3.39 *	11.57	11.59	11.21	4.52 *	15.12	15.58	15.42	6.22 *	19.66
Primclass_Uma	18.71	19.69	7.03 *	19.61	18.13	17.9	6.76 *	18.8	17.58	17.27	6.84 *	18.36
Primclass_Umar	19.8	20.59	8.3 *	22.26	15.31	14.96	6.84 *	17.09	26.64	24.42	9.6 *	28.61
Univdorm_Una	16.6	18.33	5.3 *	21.54	15.87	15.33	4.99 *	17.96	18.1	16.71	5.52 *	18.51
Primclass_Ulysses	26.55	27.09	11.97 *	28.72	23.11	22.96	10.66 *	22.05	31.51	28.85	15.31 *	29.78

*: best value.

**Figure 7.** Cont.

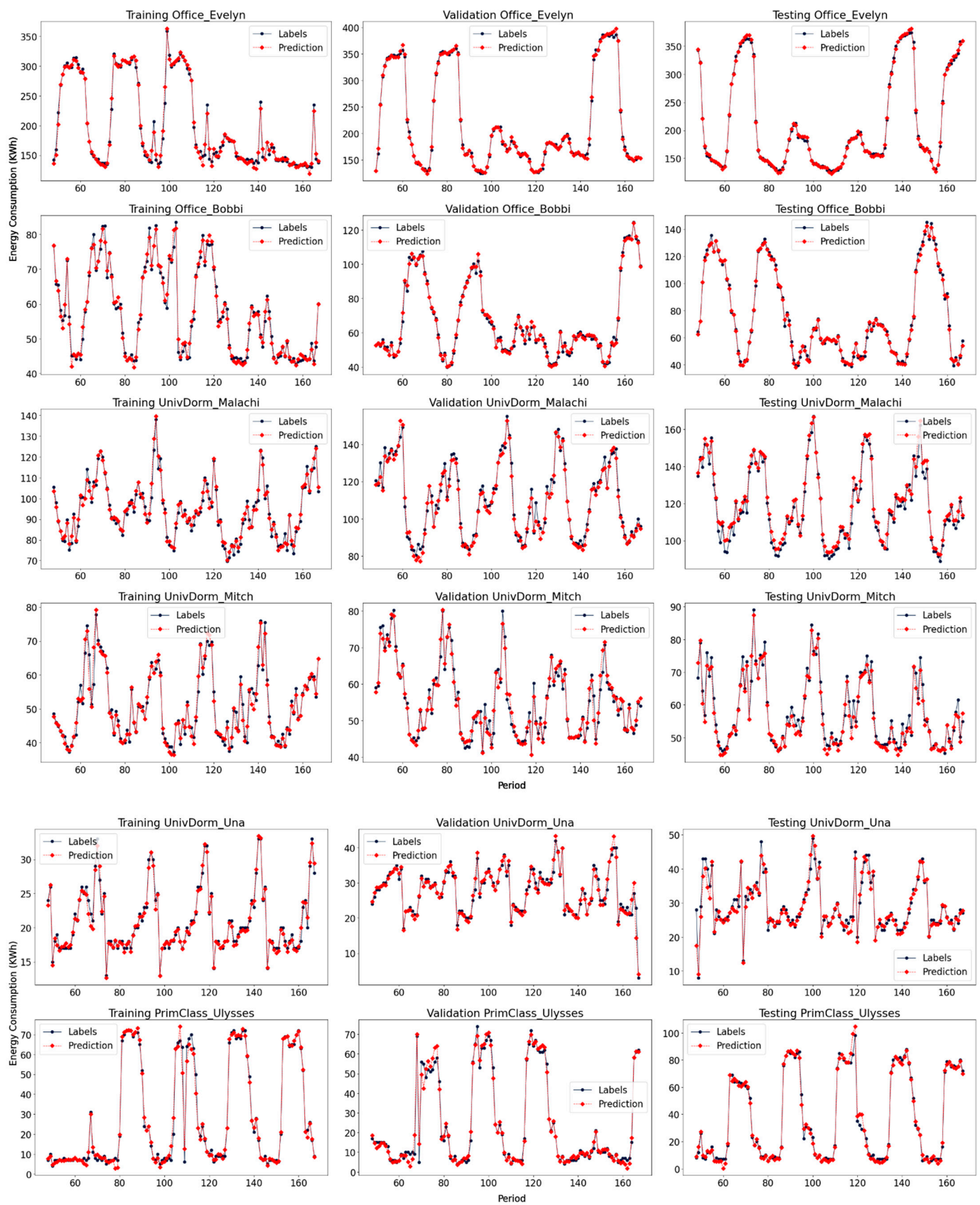


Figure 7. Cont.

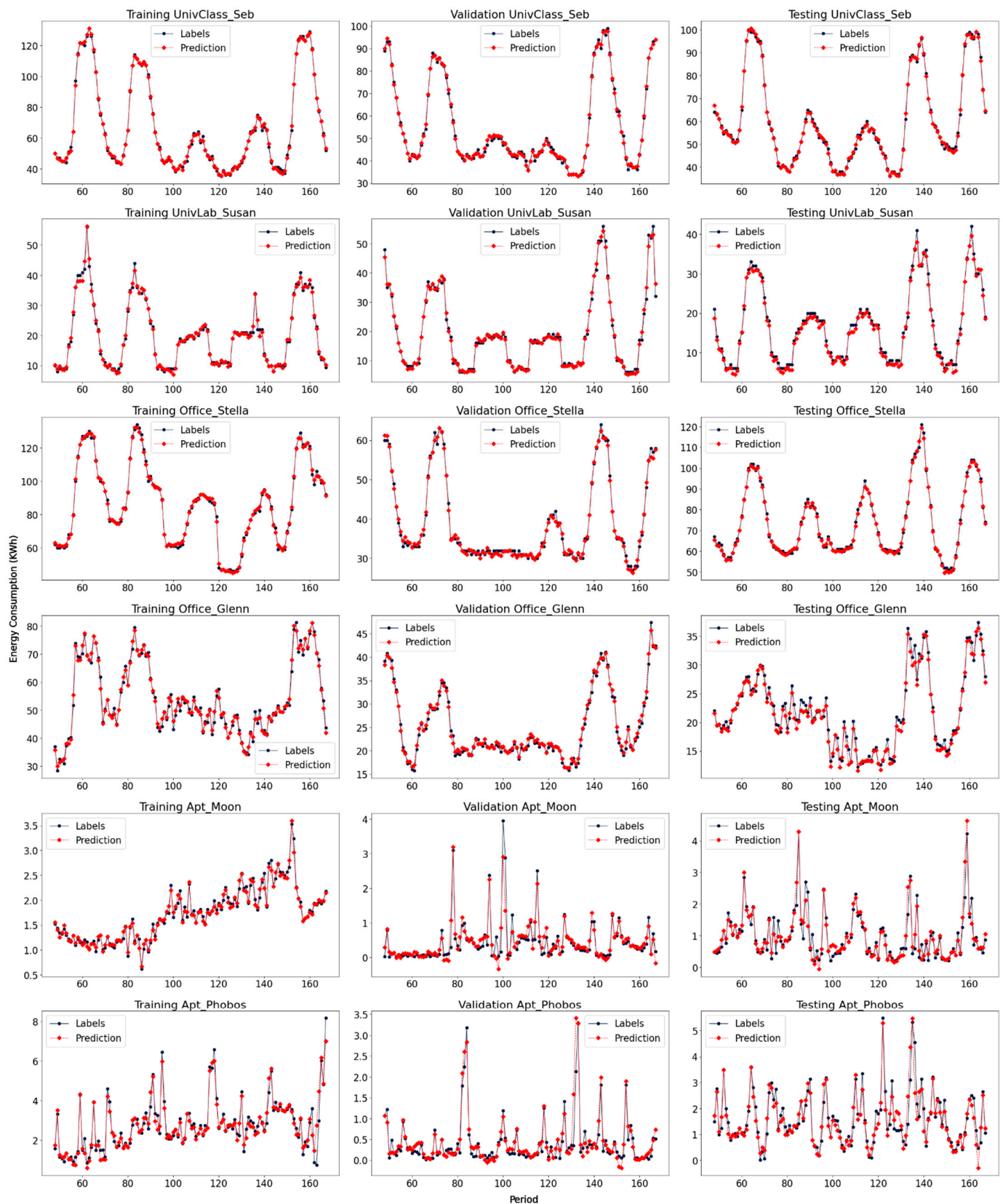


Figure 7. Cont.

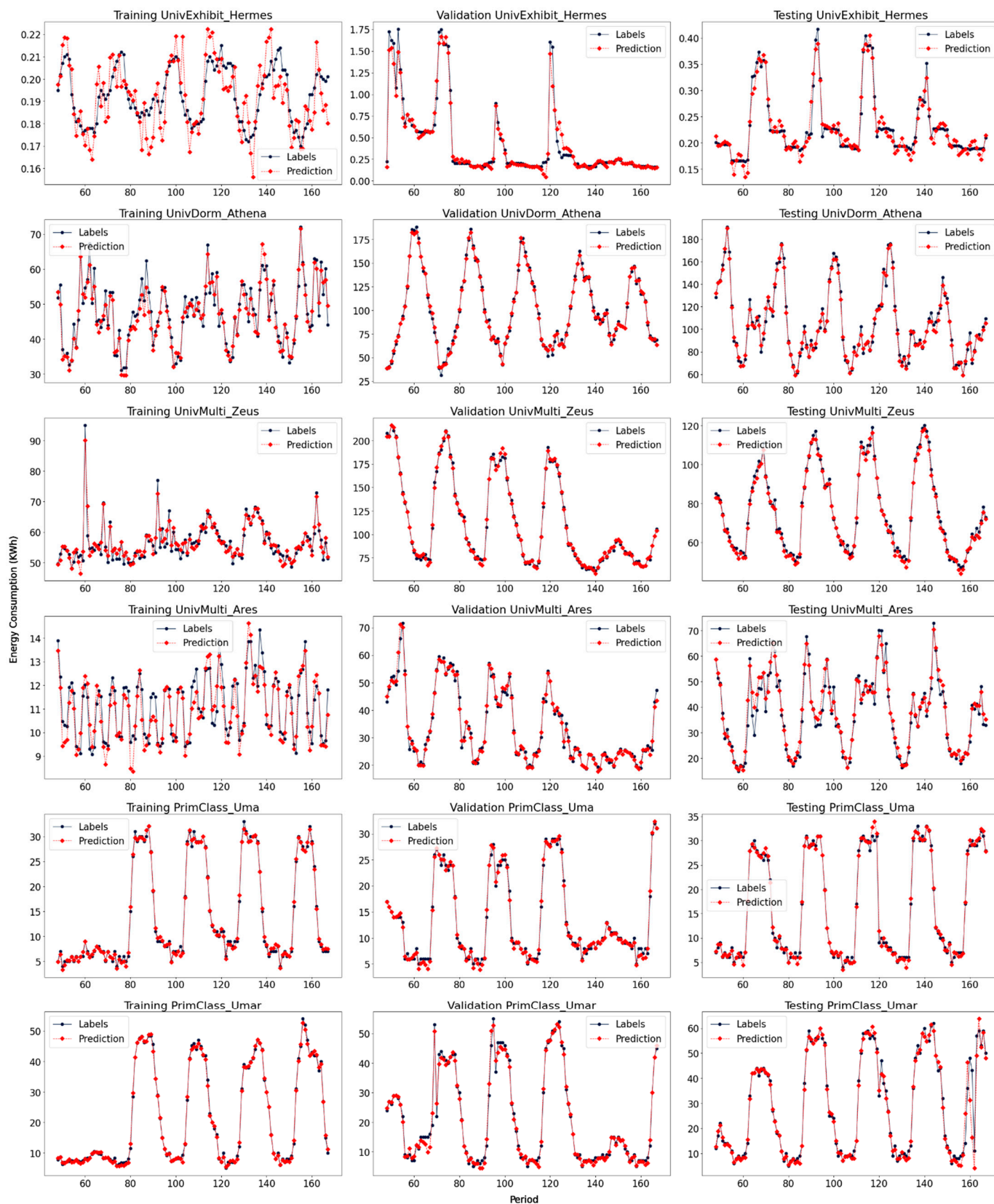


Figure 7. Comparison of the Actual and Predicted Value.

Furthermore, this study also applies a time-series cross-validation to evaluate the forecasting result. Figure 8 shows the framework of the time-series cross-validation applied

in this study. Tables 12–15 summarize the results of the time-series cross-validation using five-fold cross-validation. In addition, Table 16 shows the comparison of all algorithms. These tables show that the EEMD-LSTM algorithm obtains the best result for most of the buildings in terms of the average of MAPE, MSE, and MAE.

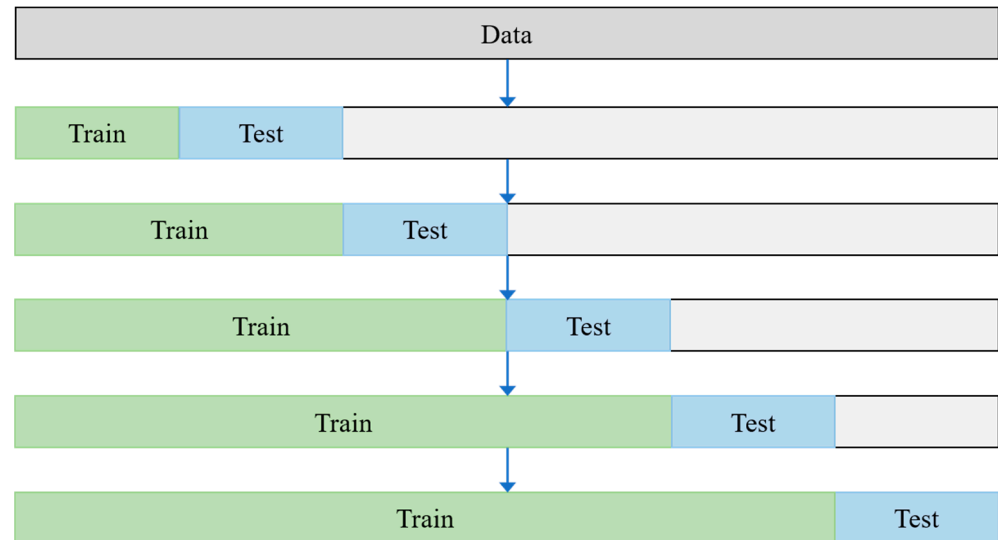


Figure 8. Framework for the time-series cross-validation.

The EEMD decomposition is carried out to the maximum decomposition based on the amount of data used on its implementation. Meanwhile, the LSTM method has the lower error rates than the WT-LSTM because the wavelet object selection uses Daubechies 1 (db1). The db1 is the same as the oldest wavelet object, namely Haar. This object is the most commonly used object to simplify the implementation, and db1 only has two filters. Each filter is used to catch the low pass filter and high pass filter. The low pass filter is used to decompose the signal that is close to the original data. It is called the approximation series. The high pass filter contains noise named as detail series. The decomposition level used single or one level. Therefore, it only produces an approximation series and a detailed series which causes the signal decomposition to be less smooth and results in a fairly high error compared to the LSTM.

Based on the forecasting results, the apartments buildings consisting of Apt_Moon and Apt_Phobos have the highest error percentage rates in training, validation, and testing. This is caused by the raw data of the two buildings having an extreme slope, touching below 1 kWh. Those two buildings have the lowest level of consumption among the other two buildings in the same category. The uneven distribution of the raw data causes high errors during the forecasting process. Both buildings have significantly different data ranges in several periods. For several periods, the data are relatively stable, while at other periods significantly decreases. The forecasting model is generated based on the training set, which is relatively stable data, while the validation and testing set has a significantly different pattern. Therefore, the validation and testing errors are significantly higher than the training error. The result of five-fold cross-validation also shows that the average MAPE of the third fold has the highest value. At the third fold, the training and test set has a significantly different pattern as illustrated in Figure 9.

Table 12. The cross-validation result of forecasting using LSTM without decomposition.

Dataset	MAPE (%)				MAE			MSE		
	Metrics	Train	Val	Test	Training	Metrics	Train	Val	Test	Testing
Univclass_Boyd	mean	1.98	1.98	1.99	0.45	0.46	0.44	0.34	0.35	0.33
	std	0.07	0.13	0.13	0.01	0.04	0.04	0.02	0.05	0.06
Univlab_Bethany	mean	2.96	3.06	3.08	2.24	2.32	2.28	10.52	10.84	10.39
	std	0.23	0.16	0.29	0.17	0.18	0.30	1.74	2.17	2.75
Office_Evelyn	mean	5.63	4.80	5.37	12.82	10.55	11.57	669.14	437.24	468.07
	std	1.30	2.40	1.68	3.55	5.98	3.80	296.19	523.21	320.99
Office_Bobbi	mean	5.79	6.87	7.22	4.34	5.04	5.07	38.45	48.06	47.68
	std	0.13	0.91	0.81	0.16	0.71	0.53	3.48	13.67	8.96
Univdorm_Malachi	mean	4.86	5.29	5.88	3.55	4.63	5.91	30.76	45.01	72.17
	std	0.52	0.80	1.18	0.68	1.72	2.15	9.96	27.51	48.69
Univdorm_Mitch	mean	8.13	8.09	8.39	5.50	5.58	5.43	60.91	58.41	56.46
	std	0.61	1.08	0.92	0.24	0.99	1.08	3.79	17.42	21.47
Univclass_Seb	mean	4.28	4.53	4.64	2.55	2.72	2.71	11.97	14.12	13.77
	std	0.82	0.68	0.64	0.44	0.58	0.41	4.46	6.71	4.08
Univlab_Susan	mean	9.63	12.22	12.63	2.08	2.50	2.38	13.45	18.08	16.36
	std	2.95	2.93	2.30	0.49	0.73	0.59	5.24	10.41	8.70
Office_Stella	mean	3.89	4.66	5.73	2.66	2.72	3.19	15.49	14.90	19.74
	std	0.52	1.25	2.37	0.42	0.55	0.59	4.56	5.22	5.80
Office_Glenn	mean	7.19	7.30	8.49	3.19	2.54	2.47	20.13	13.62	12.25
	std	0.66	1.02	1.76	0.49	1.05	0.82	4.86	10.11	8.00
Apt_Moon	mean	28.05	82.00	110.99	0.23	0.31	0.38	0.15	0.32	0.43
	std	16.15	58.70	79.86	0.01	0.08	0.11	0.05	0.19	0.23
Apt_Phobos	mean	42.10	78.63	130.29	0.50	0.45	0.53	0.71	0.60	0.65
	std	17.11	48.48	117.08	0.09	0.16	0.16	0.16	0.33	0.30
Univexhibit_Hermes	mean	5.45	8.01	9.19	0.02	0.03	0.05	0.01	0.01	0.02
	std	1.34	2.83	4.45	0.01	0.03	0.04	0.01	0.01	0.02
Univdorm_Athena	mean	8.77	9.75	10.53	8.41	10.52	11.53	121.42	184.66	255.98
	std	1.10	2.13	1.29	0.44	1.22	2.94	11.49	36.83	165.99
Univmulti_Zeus	mean	5.77	5.63	6.23	5.24	5.39	5.59	57.62	58.09	67.20
	std	0.57	0.52	1.13	0.25	0.56	0.53	6.04	11.55	15.04
Univmulti_Ares	mean	7.59	10.44	12.37	1.62	3.22	3.99	7.68	25.13	35.01
	std	0.83	2.34	2.29	0.38	1.36	1.25	3.36	17.99	17.34
Primclass_Uma	mean	20.35	24.15	20.05	1.74	2.05	1.72	11.91	16.02	10.39
	std	1.92	12.26	3.63	0.27	0.98	0.38	2.76	14.37	4.17
Primclass_Umar	mean	19.94	23.81	22.26	2.77	3.56	3.38	29.92	45.42	36.31
	std	2.05	10.09	5.26	0.27	1.20	0.47	4.99	31.71	9.52
Univdorm_Una	mean	15.99	24.26	21.37	2.32	3.46	3.45	15.69	30.04	26.86
	std	3.19	12.87	7.79	0.27	0.78	0.49	3.52	15.26	8.09
Primclass_Ulysses	mean	28.54	32.62	30.10	3.95	4.90	4.75	79.57	103.86	88.49
	std	2.34	16.08	6.46	0.45	1.84	0.80	8.32	62.30	21.51

Table 13. The cross-validation result of forecasting using RNN without decomposition.

Dataset	MAPE (%)				MAE			MSE		
	Metrics	Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	1.77	2.15	2.19	0.40	0.49	0.49	0.29	0.40	0.40
	std	0.07	0.14	0.19	0.02	0.04	0.06	0.02	0.07	0.09
Univlab_Bethany	mean	2.89	3.16	3.25	2.17	2.38	2.40	9.93	11.15	11.05
	std	0.35	0.33	0.45	0.23	0.31	0.41	1.98	3.22	3.55
Office_Evelyn	mean	5.47	5.29	5.70	11.98	11.16	11.90	543.37	413.15	453.10
	std	1.19	2.95	2.08	2.97	6.88	4.49	198.38	495.12	346.28
Office_Bobbi	mean	5.56	7.09	7.43	4.14	5.17	5.22	35.50	49.38	49.74
	std	0.20	0.91	0.69	0.14	0.78	0.50	2.45	14.41	8.89
Univdorm_Malachi	mean	5.68	6.07	6.65	3.87	5.01	6.46	31.97	47.91	84.84
	std	1.54	2.03	1.75	1.01	1.46	2.29	11.43	24.32	65.81
Univdorm_Mitch	mean	7.81	9.26	9.50	5.24	6.33	6.07	55.59	70.74	67.28
	std	0.71	1.35	1.07	0.41	1.26	1.15	5.36	23.63	24.65
Univclass_Seb	mean	4.40	4.94	5.05	2.64	2.97	2.97	13.01	16.84	16.74
	std	0.82	0.78	1.02	0.51	0.74	0.63	5.50	9.19	7.23
Univlab_Susan	mean	9.90	12.93	13.14	1.96	2.47	2.31	10.95	16.45	14.57
	std	2.87	3.40	3.41	0.33	0.68	0.58	2.39	8.78	7.26
Office_Stella	mean	3.92	5.04	6.28	2.63	2.86	3.36	13.88	14.99	20.39
	std	0.72	1.68	3.36	0.45	0.74	0.85	3.94	6.17	7.30
Office_Glenn	mean	7.11	8.38	9.45	3.08	2.80	2.73	18.68	15.37	14.02
	std	0.95	2.06	2.11	0.41	0.97	0.83	3.58	9.64	8.07
Apt_Moon	mean	34.71	106.15	130.17	0.22	0.34	0.40	0.15	0.33	0.44
	std	28.48	78.76	104.87	0.03	0.09	0.11	0.05	0.19	0.22
Apt_Phobos	mean	47.12	116.85	163.98	0.45	0.49	0.57	0.60	0.62	0.67
	std	25.02	86.25	137.74	0.05	0.17	0.15	0.09	0.38	0.31
Univexhibit_Hermes	mean	5.91	8.42	9.81	0.02	0.04	0.05	0.01	0.01	0.02
	std	1.75	2.84	4.80	0.01	0.03	0.05	0.01	0.01	0.03
Univdorm_Athena	mean	8.70	9.45	10.87	8.21	10.40	12.28	118.95	178.36	312.43
	std	1.44	1.47	2.04	1.05	0.85	4.13	28.96	24.06	274.21
Univmulti_Zeus	mean	5.99	6.06	6.69	5.38	5.78	6.13	58.70	65.71	82.24
	std	0.78	0.63	0.96	0.52	0.68	0.96	9.33	12.72	28.88
Univmulti_Ares	mean	7.82	10.74	12.56	1.62	3.33	4.05	7.40	26.69	36.11
	std	1.56	2.53	2.39	0.48	1.46	1.30	3.66	19.51	17.74
Primclass_Uma	mean	20.94	27.52	21.81	1.77	2.23	1.84	11.34	16.59	10.79
	std	2.64	14.96	4.16	0.24	1.01	0.37	2.09	14.08	3.30
Primclass_Umar	mean	21.26	27.60	25.24	2.78	3.83	3.56	27.92	49.36	37.78
	std	3.59	14.09	6.78	0.32	1.37	0.47	4.66	35.83	9.91
Univdorm_Una	mean	14.31	25.33	20.67	1.98	3.46	3.37	12.10	29.67	25.06
	std	3.51	14.22	5.94	0.32	0.82	0.19	3.46	16.16	2.29
Primclass_Ulysses	mean	33.94	42.54	37.12	4.05	5.48	5.11	77.71	116.37	90.65
	std	6.53	26.52	11.83	0.67	2.33	1.03	15.01	83.52	27.24

Table 14. The cross-validation result of forecasting using EEMD-LSTM.

Dataset	Metrics	MAPE (%)			MAE			MSE		
		Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	0.63	0.71	0.88	0.14	0.16	0.19	0.03	0.04	0.06
	std	0.03	0.13	0.30	0.01	0.03	0.06	0.00	0.01	0.03
Univlab_Bethany	mean	1.05	1.07	1.23	0.79	0.80	0.89	1.18	1.16	1.39
	std	0.07	0.06	0.32	0.06	0.09	0.21	0.20	0.32	0.58
Office_Evelyn	mean	2.47	2.01	2.42	5.16	4.25	4.92	85.82	69.47	64.00
	std	0.74	1.30	1.23	1.68	3.08	2.42	51.94	115.69	59.93
Office_Bobbi	mean	1.96	2.17	2.43	1.40	1.56	1.65	3.46	4.21	4.71
	std	0.08	0.18	0.24	0.07	0.23	0.18	0.35	1.22	0.85
Univdorm_Malachi	mean	2.07	1.78	5.79	1.36	1.50	5.13	3.59	4.45	47.03
	std	0.42	0.26	3.18	0.31	0.58	2.94	1.41	3.06	44.22
Univdorm_Mitch	mean	2.87	2.85	5.97	1.86	1.89	3.70	7.32	7.07	24.62
	std	0.31	0.59	2.70	0.09	0.37	2.00	0.59	2.52	21.42
Univclass_Seb	mean	1.50	1.59	2.27	0.88	0.91	1.26	1.30	1.40	2.78
	std	0.09	0.10	0.90	0.05	0.11	0.45	0.15	0.37	1.69
Univlab_Susan	mean	4.54	5.53	6.19	0.85	1.04	1.05	1.86	3.22	3.16
	std	0.62	1.18	1.15	0.09	0.33	0.25	0.25	2.24	1.94
Office_Stella	mean	1.37	2.61	11.09	0.93	1.30	4.76	1.44	3.93	50.78
	std	0.12	2.32	12.87	0.10	0.85	4.55	0.29	5.54	64.65
Office_Glenn	mean	2.31	2.34	8.09	1.00	0.77	2.35	1.94	1.20	12.21
	std	0.15	0.32	6.12	0.12	0.23	1.94	0.45	0.80	14.86
Apt_Moon	mean	17.63	59.03	77.37	0.08	0.13	0.25	0.02	0.05	0.13
	std	14.99	48.50	38.83	0.01	0.05	0.13	0.01	0.03	0.10
Apt_Phobos	mean	26.91	49.46	77.26	0.21	0.19	0.34	0.10	0.10	0.28
	std	14.49	31.34	44.27	0.05	0.09	0.29	0.04	0.09	0.39
Univexhibit_Hermes	mean	5.18	6.93	15.17	0.01	0.02	0.05	0.00	0.00	0.01
	std	0.55	1.63	8.85	0.00	0.01	0.04	0.00	0.00	0.01
Univdorm_Athena	mean	3.14	2.84	4.94	2.94	3.09	5.28	13.88	15.47	57.58
	std	0.38	0.41	2.77	0.22	0.27	3.44	2.04	2.52	75.26
Univmulti_Zeus	mean	2.06	2.14	3.90	1.79	1.93	3.33	5.83	6.34	22.32
	std	0.25	0.44	2.18	0.12	0.19	2.03	0.79	1.11	26.55
Univmulti_Ares	mean	3.54	3.96	7.39	0.67	1.15	2.00	0.91	3.22	7.38
	std	0.42	0.78	3.80	0.15	0.48	0.43	0.46	2.53	2.59
Primclass_Uma	mean	7.74	11.87	10.41	0.72	1.00	0.88	1.82	3.83	2.32
	std	1.01	9.63	4.44	0.10	0.66	0.26	0.40	4.30	1.42
Primclass_Umar	mean	8.32	11.38	10.39	1.08	1.57	1.43	3.89	9.90	6.80
	std	0.90	6.47	3.01	0.10	0.72	0.27	0.76	9.33	2.96
Univdorm_Una	mean	5.21	9.43	7.85	0.92	1.48	1.48	2.44	6.41	5.02
	std	0.68	8.07	3.61	0.09	0.69	0.39	0.55	6.02	1.85
Primclass_Ulysses	mean	13.42	17.74	18.04	1.64	2.29	2.14	9.61	20.13	13.61
	std	1.49	11.35	7.26	0.20	1.26	0.64	1.61	19.11	6.44

Table 15. The cross-validation result of forecasting using WT-LSTM.

Dataset	Metrics	MAPE (%)			MAE			MSE		
		Train	Val	Test	Train	Val	Test	Train	Val	Test
Univclass_Boyd	mean	2.30	2.24	2.27	0.52	0.52	0.51	0.46	0.46	0.44
	std	0.06	0.14	0.15	0.01	0.04	0.05	0.02	0.06	0.08
Univlab_Bethany	mean	4.04	4.07	4.31	3.06	3.12	3.17	20.75	21.13	20.06
	std	0.42	0.24	0.45	0.34	0.35	0.41	5.21	6.86	4.97
Office_Evelyn	mean	8.64	7.17	8.50	19.34	15.36	17.99	1159.95	742.98	961.04
	std	1.57	2.00	1.90	4.38	5.63	4.44	421.95	513.37	404.27
Office_Bobbi	mean	7.76	8.37	8.94	5.73	6.11	6.18	62.40	71.40	70.87
	std	0.26	1.13	0.95	0.21	1.05	0.69	4.88	24.61	16.28
Univdorm_Malachi	mean	5.88	5.83	6.54	4.16	5.10	6.57	38.91	53.32	87.12
	std	0.39	0.66	1.41	0.61	1.98	2.52	11.40	32.06	58.59
Univdorm_Mitch	mean	9.70	8.50	9.21	6.27	5.74	5.80	72.26	58.66	61.10
	std	1.13	1.38	1.69	0.24	0.95	1.32	4.71	15.52	24.02
Univclass_Seb	mean	5.60	6.12	6.47	3.35	3.67	3.77	23.80	29.28	29.66
	std	0.99	1.06	0.85	0.63	0.73	0.34	8.79	13.46	6.30
Univlab_Susan	mean	12.18	14.85	15.93	2.53	3.00	2.90	16.77	23.83	21.88
	std	2.00	3.73	2.79	0.30	0.83	0.54	2.49	12.17	10.01
Office_Stella	mean	5.79	6.06	7.41	3.96	3.64	4.36	31.26	26.88	36.13
	std	0.57	0.58	1.35	0.58	0.74	0.63	8.41	9.87	10.53
Office_Glenn	mean	8.39	8.13	9.23	3.77	2.87	2.80	28.07	18.35	16.85
	std	0.31	1.05	1.09	0.49	1.13	0.92	5.63	12.81	11.52
Apt_Moon	mean	37.94	114.79	131.96	0.26	0.33	0.41	0.18	0.34	0.46
	std	27.42	89.40	95.90	0.02	0.10	0.12	0.04	0.21	0.25
Apt_Phobos	mean	58.37	100.42	134.13	0.60	0.47	0.58	0.92	0.62	0.69
	std	22.89	58.93	63.67	0.14	0.19	0.21	0.30	0.40	0.34
Univexhibit_Hermes	mean	8.43	12.06	13.58	0.03	0.05	0.06	0.01	0.02	0.03
	std	2.17	5.10	6.91	0.02	0.03	0.05	0.01	0.02	0.03
Univdorm_Athena	mean	11.68	11.10	13.30	11.12	12.07	14.95	210.67	245.96	432.71
	std	2.33	2.11	1.94	2.28	1.15	4.76	91.05	43.07	312.91
Univmulti_Zeus	mean	6.70	6.49	7.29	6.13	6.38	6.80	77.49	89.44	106.63
	std	0.61	0.38	0.72	0.25	0.82	0.87	3.99	24.35	26.57
Univmulti_Ares	mean	12.62	14.36	16.27	2.52	4.35	5.22	15.05	42.18	57.08
	std	0.65	2.45	2.16	0.51	1.69	1.47	6.16	28.26	26.77
Primclass_Uma	mean	21.32	24.95	21.78	2.06	2.21	2.04	15.09	17.83	13.81
	std	2.69	13.29	2.99	0.34	0.98	0.37	3.30	13.89	4.91
Primclass_Umar	mean	23.74	25.86	26.54	3.50	3.96	4.15	37.80	45.52	45.54
	std	4.30	9.47	5.03	0.49	0.99	0.32	7.91	23.52	7.67
Univdorm_Una	mean	17.37	22.50	23.06	2.59	3.37	3.94	16.29	26.25	30.73
	std	4.03	9.72	6.84	0.34	0.61	0.89	4.34	12.45	12.21
Primclass_Ulysses	mean	34.30	37.13	34.27	5.02	5.57	5.52	103.54	118.14	109.34
	std	4.40	17.17	7.33	0.70	1.72	0.81	16.75	57.71	30.24

Table 16. Comparison of algorithms based on MAPE (%) using time series cross-validation.

Dataset	Training				Validation				Testing			
	RNN	LSTM	EEMD LSTM	WT LSTM	RNN	LSTM	EEMD LSTM	WT LSTM	RNN	LSTM	EEMD LSTM	WT LSTM
Univclass_Boyd	1.77	1.98	0.63 *	2.3	2.15	1.98	0.71 *	2.24	2.19	1.99	0.88 *	2.27
Univlab_Bethany	2.89	2.96	1.05 *	4.04	3.16	3.06	1.07 *	4.07	3.25	3.08	1.23 *	4.31
Office_Evelyn	5.47	5.63	2.47 *	8.64	5.29	4.8	2.01 *	7.17	5.7	5.37	2.42 *	8.5
Office_Bobbi	5.56	5.79	1.96 *	7.76	7.09	6.87	2.17 *	8.37	7.43	7.22	2.43 *	8.94
Univdorm_Malachi	5.68	4.86	2.07 *	5.88	6.07	5.29	1.78 *	5.83	6.65	5.88	5.79 *	6.54
Univdorm_Mitch	7.81	8.13	2.87 *	9.7	9.26	8.09	2.85 *	8.5	9.5	8.39	5.97 *	9.21
Univclass_Seb	4.4	4.28	1.5 *	5.6	4.94	4.53	1.59 *	6.12	5.05	4.64	2.27 *	6.47
Univlab_Susan	9.9	9.63	4.54 *	12.18	12.93	12.22	5.53 *	14.85	13.14	12.63	6.19 *	15.93
Office_Stella	3.92	3.89	1.37 *	5.79	5.04	4.66	2.61 *	6.06	6.28	5.73 *	11.09	7.41
Office_Glenn	7.11	7.19	2.31 *	8.39	8.38	7.3	2.34 *	8.13	9.45	8.49	8.09 *	9.23
Apt_Moon	34.71	28.05	17.63 *	37.94	106.15	82	59.03 *	114.79	130.2	110.99	77.37 *	131.96
Apt_Phobos	47.12	42.1	26.91 *	58.37	116.85	78.63	49.46 *	100.42	163.9	130.29	77.26 *	134.13
Univexhibit_Hermes	5.91	5.45	5.18 *	8.43	8.42	8.01	6.93 *	12.06	9.81	9.19 *	15.17	13.58
Univdorm_Athena	8.7	8.77	3.14 *	11.68	9.45	9.75	2.84 *	11.1	10.87	10.53	4.94 *	13.3
Univmulti_Zeus	5.99	5.77	2.06 *	6.7	6.06	5.63	2.14 *	6.49	6.69	6.23	3.9 *	7.29
Univmulti_Ares	7.82	7.59	3.54 *	12.62	10.74	10.44	3.96 *	14.36	12.56	12.37	7.39 *	16.27
Primclass_Uma	20.94	20.35	7.74 *	21.32	27.52	24.15	11.87 *	24.95	21.81	20.05	10.41 *	21.78
Primclass_Umar	21.26	19.94	8.32 *	23.74	27.6	23.81	11.38 *	25.86	25.24	22.26	10.39 *	26.54
Univdorm_Una	14.31	15.99	5.21 *	17.37	25.33	24.26	9.43 *	22.5	20.67	21.37	7.85 *	23.06
Primclass_Ulysses	33.94	28.54	13.42 *	34.3	42.54	32.62	17.74 *	37.13	37.12	30.1	18.04 *	34.27

*: best value

Buildings such as college classrooms and laboratories have more stable forecasting results with small errors than other buildings, followed by offices. Another building that has a high error is the primary/secondary classroom. It is caused by the fairly high distribution of data based on the buildings in that category. Stationary buildings such as UnivClass_Boyd, Office_Stella and PrimClass_Uma tend to have small and stable errors based on data stationarity. The UnivClass_Boyd has the smallest error among other buildings. Likewise, UnivDorm_Malachi is the only non-stationary building that has very small and stable errors in training, validation, and testing.

Further comparison is conducted using statistics Levene's test and T-test to evaluate how significant the differences are between the LSTM without decomposition and with decomposition. The statistics test uses MAPE results as the data. This study runs the statistical tests using SPSS.

Levene's test checks whether the variance of the two samples is the same or not with the following hypothesis below.

H_0 —both samples have the same variance

H_1 —both samples have different variances

If the p -value (sig.) is less than 0.05, then reject H_0 meaning that the results from the two methods have different variances. Otherwise, they have the same variances.

The T-test compares the results based on their means with the following hypothesis.

H_0 —both samples have the similar means

H_1 —both samples have different means

If the p -value (sig.) is less than 0.05, then reject H_0 meaning that the results from the two methods are significantly different. Otherwise, their results are not significantly different.

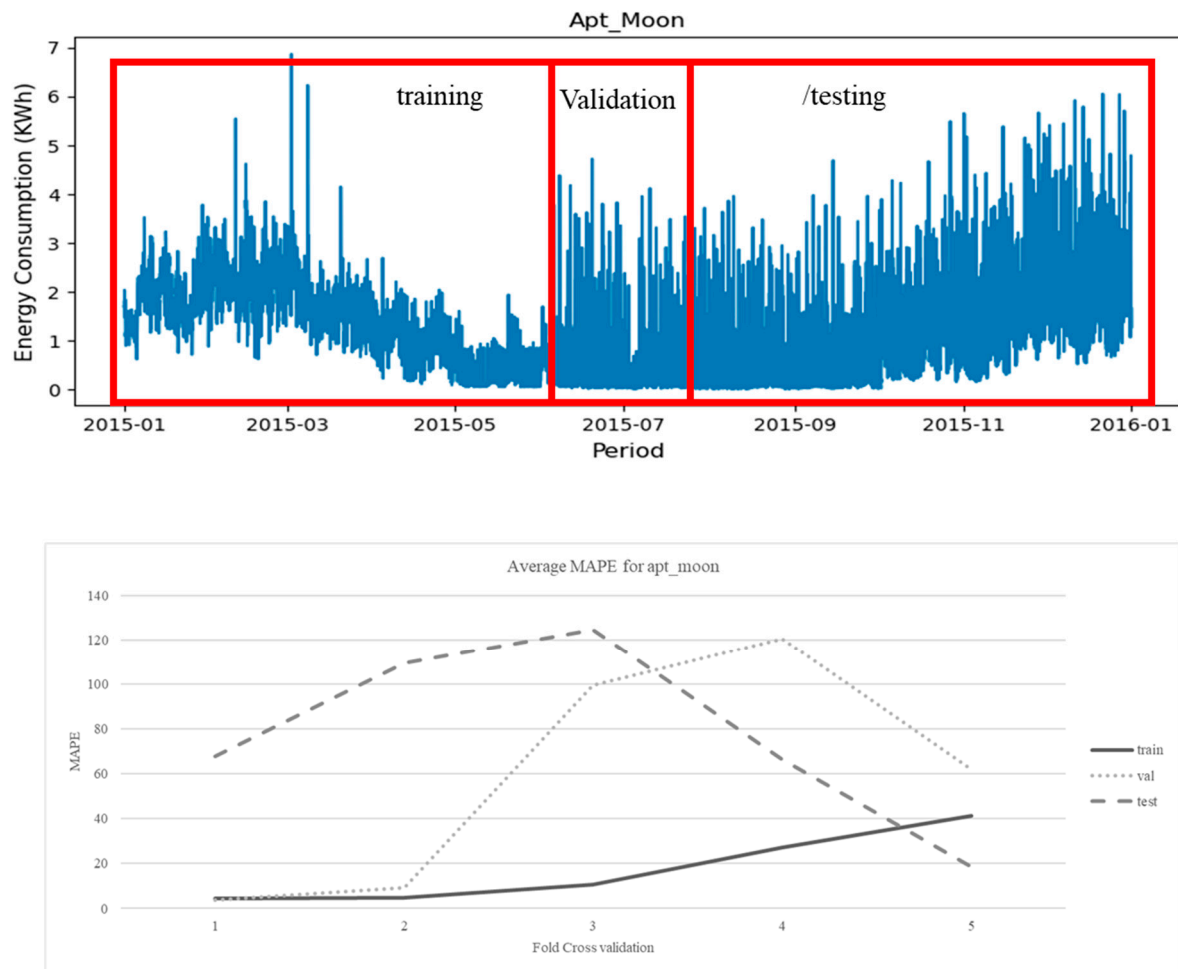


Figure 9. The pattern differences between training, validation, and testing set for Apt_moon dataset.

The statistics test results summarized in Table 17 reveal that the LSTM and EEMD-LSTM methods have significantly different variance and average. However, LSTM and WT-LSTM do not significantly differ in terms of the variance and average. Furthermore, the results of EEMD-LSTM and WT-LSTM are significantly different in terms of variance and average.

Table 17. Statistics' test results.

Comparison	Mean	Std. Deviation	p -Value of Levene's Test	p -Value of T-Test
LSTM	0.153	0.145	-	
EEMD-LSTM	0.076	0.062		
WT-LSTM	0.164	0.135		
LSTM vs. EEMD-LSTM	-	-	0.008 Reject H_0	0.037 Reject H_0
LSTM vs. WT-LSTM	-	-	0.852 Accept H_0	0.813 Accept H_0
EEMD-LSTM vs. WT-LSTM	-	-	0.008 Reject H_0	0.013 Reject H_0

The test results conclude that the EEMD-LSTM method has a significant difference compared with the LSTM and WT-LSTM methods. On the other hand, the LSTM method is not significantly different from the WT-LSTM method. This conclusion is also supported by the comparison of the average MAPE Test for each method, where the MAPE test for the EEMD-LSTM method tends to be lower. However, the MAPE of LSTM and WT-LSTM methods are not significantly different. The results also reveal that the EEMD-LSTM method obtains better results than the LSTM and WT-LSTM methods.

Based on the previous studies conducted by Imani and Ghassemian [49] and Fang, Gao, and Ruan [14], the error of LSTM without decomposition is between 3% to 16%. Meanwhile, for this research, the forecasting results using the LSTM method has an average error of 15.3%. The results show that EEMD has successfully improved the forecasting accuracy of the LSTM method.

Finally, this research also provides the recommendation of the forecasting method for each building, as listed in Table 18. However, for the APT-Moon and APT_Phobos, other algorithms should be considered because the MAPE of the EEMD-LSTM is still above 20%.

Table 18. Recommendation of the forecasting method.

Units	Functionalities	Time Zones	Recommendation
Office_Bobbi	Office	America/Los_Angeles	EEMD-LSTM
Office_Evelyn	Office	America/Los_Angeles	EEMD-LSTM
UnivClass_Boyd	College Classroom	America/Los_Angeles	EEMD-LSTM
UnivLab_Bethany	College Laboratory	America/Los_Angeles	EEMD-LSTM
UnivDorm_Malachi	Residential	America/Chicago	EEMD-LSTM
UnivDorm_Mitch	Residential	America/Chicago	EEMD-LSTM
Office_Glenn	Office	Europe/London	EEMD-LSTM
Office_Stella	Office	Europe/London	EEMD-LSTM
UnivClass_Seb	College Classroom	Europe/London	EEMD-LSTM
UnivLab_Susan	College Laboratory	Europe/London	EEMD-LSTM
Apt_Moon	Apartment	America/New_York	EEMD-LSTM
Apt_Phobos	Apartment	America/New_York	EEMD-LSTM
UnivDorm_Athena	Residential	Asia/Taiwan	EEMD-LSTM
UnivExhibit_Hermes	Exhibition	Asia/Taiwan	EEMD-LSTM
UnivMulti_Ares	Multipurpose	Asia/Taiwan	EEMD-LSTM
UnivMulti_Zeus	Multipurpose	Asia/Taiwan	EEMD-LSTM
PrimClass_Ulysses	Secondary Classroom	Asia/Singapore	EEMD-LSTM
PrimClass_Uma	Secondary Classroom	Asia/Singapore	EEMD-LSTM
PrimClass_Umar	Secondary Classroom	Asia/Singapore	EEMD-LSTM
UnivDorm_Una	Residential	Asia/Singapore	EEMD-LSTM

5. Conclusions

Energy consumption is predicted to increase continuously. Building as a facility where most of the human activities happen consumes energy, especially electricity. Without proper energy management, energy consumption of a building might be inefficient. On the other hand, managing energy consumption needs information on the future energy demand. Therefore, this study proposes a forecasting framework to predict the energy consumption of a building.

The proposed forecasting framework is developed based on LSTM with decomposition methods to improve the performance of the LSTM. There are three methods evaluated

in this study. They are the LSTM without decomposition, LSTM with EEMD as the decomposition, and the LSTM with WT as the decomposition method. The LSTM without decomposition is the baseline for both EEMD-LSTM and WT-LSTM.

This study applies the proposed forecasting methods on 20 series of energy consumption data from 20 different buildings located in many different countries and different functions. The experiment shows that the EEMD-LSTM method obtains the smallest error for 85% of the series. While 15% of the series obtained the smallest error when using the LSTM method. The average error for LSTM, EEMD-LSTM and WT-LSTM are 15.31%, 7.57%, and 16.36%, respectively. The statistics T-test also reveals that the EEMD-LSTM method is significantly better than the other two methods.

Meanwhile, the results of the LSTM and WT-LSTM are not significantly different. The EEMD and WT methods have significantly different values because the EEMD decomposed the data based on the sample data used. On the other hand, the WT decomposition only uses a single-level decomposition, and the object wavelets only have two filters. Further study should increase the decomposition level and the number of filters to improve the performance of WT. Other hybrid forecasting methods and differences also should be considered. In addition to the forecasting algorithm, further research should consider more datasets from different types and locations of the building to get a more general forecasting framework.

Author Contributions: Conceptualization, A.D. and S.-Y.C.; methodology, F.E.Z.; software, A.D. and M.F.; validation, A.D. and F.E.Z.; formal analysis, A.D. and F.E.Z.; investigation, A.D., S.-Y.C. and F.E.Z.; resources, S.-Y.C.; data curation, A.D.; writing, A.D., F.E.Z. and M.F.; visualization, M.F.; supervision, S.-Y.C.; project administration, S.-Y.C.; funding acquisition, S.-Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded in part by the Taiwan Building Technology Center, National Taiwan University of Science and Technology, from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan under the grant 110P011.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Benchmark data are available on <https://github.com/buds-lab/the-building-data-genome-project> [44] (accessed on 13 March 2021) and <http://traces.cs.umass.edu/index.php/Smart/Smart> [45] (accessed on 13 March 2021).

Acknowledgments: This work was supported in part by the Taiwan Building Technology Center, National Taiwan University of Science and Technology, from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan under the grant 110P011.

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

References

1. Cao, X.; Dai, X.; Liu, J. Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy Build.* **2016**, *128*, 198–213. [CrossRef]
2. Saidur, R.; Masjuki, H.H.; Jamaluddin, M.Y. An application of energy and exergy analysis in residential sector of Malaysia. *Energy Policy* **2007**, *35*, 1050–1063. [CrossRef]
3. Allouhi, A.; El Fouih, Y.; Kousksou, T.; Jamil, A.; Zeraoui, Y.; Mourad, Y. Energy consumption and efficiency in buildings: Current status and future trends. *J. Clean. Prod.* **2015**, *109*, 118–130. [CrossRef]
4. EIA. *International Energy Outlook 2019*; U.S. Energy Information Administration (EIA): Washington, DC, USA, 2019.
5. Bedi, J.; Toshniwal, D. Empirical mode decomposition based deep learning for electricity demand forecasting. *IEEE Access* **2018**, *6*, 49144–49156. [CrossRef]
6. Ahmad, T.; Chen, H. A review on machine learning forecasting growth trends and their real-time applications in different energy systems. *Sustain. Cities Soc.* **2020**, *54*, 102010. [CrossRef]
7. Kontokosta, C.E.; Tull, C. A data-driven predictive model of city-scale energy use in buildings. *Appl. Energy* **2017**, *197*, 303–317. [CrossRef]

8. Hong, W.-C.; Dong, Y.; Zhang, W.Y.; Chen, L.-Y.; Panigrahi, B.K. Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm. *Int. J. Electr. Power Energy Syst.* **2013**, *44*, 604–614. [\[CrossRef\]](#)
9. Kneifel, J.; Webb, D. Predicting energy performance of a net-zero energy building: A statistical approach. *Appl. Energy* **2016**, *178*, 468–483. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Barhmi, S.; Elfatni, O.; Belhaj, I. Forecasting of wind speed using multiple linear regression and artificial neural networks. *Energy Syst.* **2019**, *11*, 935–946. [\[CrossRef\]](#)
11. Montgomery, D.; Jennings, C.; Kulahci, M. *Introduction to Time Series Analysis and Forecasting*; John Wiley & Sons: Hoboken, NJ, USA, 2008; p. 472.
12. Mocanu, E.; Nguyen, P.H.; Gibescu, M.; Kling, W.L. Deep learning for estimating building energy consumption. *Sustain. Energy Netw.* **2016**, *6*, 91–99. [\[CrossRef\]](#)
13. Kong, W.; Dong, Z.Y.; Jia, Y.; Hill, D.J.; Xu, Y.; Zhang, Y. Short-term residential load forecasting based on lstm recurrent neural network. *IEEE Trans. Smart Grid* **2019**, *10*, 841–885. [\[CrossRef\]](#)
14. Fang, C.; Gao, Y.; Ruan, Y. Improving forecasting accuracy of daily energy consumption of office building using time series analysis based on wavelet transform decomposition. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *294*, 012031. [\[CrossRef\]](#)
15. Peng, L.; Wang, L.; Xia, D.; Gao, Q.J.E. Effective energy consumption forecasting using empirical wavelet transform and long short-term memory. *Energy* **2022**, *238*, 121756. [\[CrossRef\]](#)
16. Gao, Y.; Ruan, Y.J.E. Buildings, interpretable deep learning model for building energy consumption prediction based on attention mechanism. *Energy Build.* **2021**, *252*, 111379. [\[CrossRef\]](#)
17. Somu, N.; MR, G.R.; Ramamritham, K.J.A.E. A hybrid model for building energy consumption forecasting using long short term memory networks. *Appl. Energy* **2020**, *261*, 114131. [\[CrossRef\]](#)
18. Runge, J.; Zmeureanu, R. Forecasting energy use in buildings using artificial neural networks: A review. *Energies* **2019**, *12*, 3254. [\[CrossRef\]](#)
19. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
20. Koprinska, I.; Rana, M.; Agelidis, V.G. Correlation and instance based feature selection for electricity load forecasting. *Knowl.-Based Sys.* **2015**, *82*, 29–40. [\[CrossRef\]](#)
21. Friedrich, L.; Afshari, A. Short-term forecasting of the abu dhabi electricity load using multiple weather variables. *Energy Procedia* **2015**, *75*, 3014–3026. [\[CrossRef\]](#)
22. Deb, C.; Zhang, F.; Yang, J.; Lee, S.E.; Shah, K.W. A review on time series forecasting techniques for building energy consumption. *Renew. Sustain. Energy Rev.* **2017**, *74*, 902–924. [\[CrossRef\]](#)
23. Bourdeau, M.; Zhai, X.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustain. Cities Soc.* **2019**, *48*, 101533. [\[CrossRef\]](#)
24. Mandic, D.P.; Chambers, J.A. *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*; Wiley: Hoboken, NJ, USA, 2001.
25. Kumar, D.N.; Raju, K.S.; Sathish, T. River flow forecasting using recurrent neural networks. *Water Resour. Manag.* **2004**, *18*, 143–161. [\[CrossRef\]](#)
26. Hua, Y.; Zhao, Z.; Li, R.; Chen, X.; Liu, Z.; Zhang, H. Deep learning with long short-term memory for time series prediction. *IEEE Commun Mag.* **2019**, *57*, 114–119. [\[CrossRef\]](#)
27. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [\[CrossRef\]](#)
28. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#)
29. Aungiers, J. LSTM Neural Network for Time Series Prediction. Available online: <https://www.jakob-aungiers.com/articles/a/LSTM-Neural-Network-for-Time-Series-Prediction> (accessed on 13 March 2021).
30. Wu, Y.-X.; Wu, Q.-B.; Zhu, J.-Q. Improved EEMD-based crude oil price forecasting using LSTM networks. *Phys. A Stat. Mech. its Appl.* **2019**, *516*, 114–124. [\[CrossRef\]](#)
31. Xu, Y.; Zhou, C.; Geng, J.; Gao, S.; Wang, P. A method for diagnosing mechanical faults of on-load tap changer based on ensemble empirical mode decomposition, volterra model and decision acyclic graph support vector machine. *IEEE Access* **2019**, *7*, 84803–84816. [\[CrossRef\]](#)
32. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [\[CrossRef\]](#)
33. Huang, N.; Shen, Z.; Long, S. A new view of non-linear water waves: The hilbert spectrum. *Annu. Rev. Fluid Mech.* **2003**, *31*, 417–457. [\[CrossRef\]](#)
34. Zhang, X.; Lai, K.K.; Wang, S.-Y. A new approach for crude oil price analysis based on empirical mode decomposition. *Energy Econ.* **2008**, *30*, 905–918. [\[CrossRef\]](#)
35. Terzija, N. Robust Digital Image Watermarking Algorithms for Copyright Protection. Ph.D. Thesis, University of Duisburg-Essen, Duisburg, Germany, 2006.
36. Sripathi, D. Efficient Implementations of Discrete Wavelet Transforms using FPGAs; Florida State University, Florida, USA, 2003. Available online: http://purl.flvc.org/fdu/fdu_migr_etd-1599 (accessed on 18 July 2021).

37. Sugiartawan, P.; Pulungan, R.; Kartika, A. Prediction by a hybrid of wavelet transform and long-short-term-memory neural networks. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, e0142064. [[CrossRef](#)]
38. Jin, J.; Kim, J. Forecasting natural gas prices using wavelets, time series, and artificial neural networks. *PLoS ONE* **2015**, *10*, e0142064. [[CrossRef](#)]
39. Chang, Z.; Zhang, Y.; Chen, W. Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. *Energy* **2019**, *187*, 115804. [[CrossRef](#)]
40. Liu, Y.; Guan, L.; Hou, C.; Han, H.; Liu, Z.; Sun, Y.; Zheng, M. Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform. *Appl. Sci.* **2019**, *9*, 1108. [[CrossRef](#)]
41. Wang, F.; Yu, Y.; Zhang, Z.; Li, J.; Zhen, Z.; Li, K. Wavelet decomposition and convolutional lstm networks based improved deep learning model for solar irradiance forecasting. *Appl. Sci.* **2018**, *8*, 1286. [[CrossRef](#)]
42. Qin, Q.; Lai, X.; Zou, J. Direct multistep wind speed forecasting using lstm neural network combining eemd and fuzzy entropy. *Appl. Sci.* **2019**, *9*, 126. [[CrossRef](#)]
43. Liu, W.; Liu, W.D.; Gu, J. Forecasting oil production using ensemble empirical model decomposition based long short-term memory neural network. *J. Pet. Sci. Eng.* **2020**, *189*, 107013. [[CrossRef](#)]
44. Miller, C.; Meggers, F. The building data genome project: An open, public data set from non-residential building electrical meters. *Energy Procedia* **2017**, *122*, 439–444. [[CrossRef](#)]
45. Barker, S.; Mishra, A.; Irwin, D.; Cecchet, E.; Shenoy, P.; Albrecht, J. Smart*: An open data set and tools for enabling research in sustainable homes. *SustKDD* **2012**, *111*, 108.
46. Kwiatkowski, D.; Phillips, P.C.; Schmidt, P.; Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econ.* **1992**, *54*, 159–178. [[CrossRef](#)]
47. Singh, A.A. Gentle Introduction to Handling a Non-Stationary Time Series in Python. Available online: <https://www.analyticsvidhya.com/blog/2018/09/non-stationary-time-series-python/#> (accessed on 18 July 2021).
48. Mueller, J.P.; Massaron, L. *Machine Learning for Dummies*; John Wiley & Sons: Hoboken, NJ, USA, 2021.
49. Imani, M.; Ghassemian, H. Lagged load wavelet decomposition and lstm networks for short-term load forecasting. In Proceedings of the 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran, Iran, 6–7 March 2019; pp. 6–12.