Abstract: Electricity demand forecasting enables the stable operation of electric power systems and reduces electric power consumption. Previous studies have predicted electricity demand through a correlation analysis between power consumption and weather data; however, this analysis does not consider the influence of various factors on power consumption, such as industrial activities, economic factors, power horizon, and resident living patterns of buildings. This study proposes an efficient power demand prediction using deep learning techniques for two industrial buildings with different power consumption patterns. The problems are presented by analyzing the correlation between the power consumption and weather data by season for industrial buildings with different power consumption patterns. Four models were analyzed using the most important factors for predicting power consumption and weather data (temperature, humidity, sunlight, solar radiation, total cloud cover, wind speed, wind direction, humidity, and vapor pressure). The prediction horizon for power consumption forecasting was kept at 24 h. The existing deep learning methods (DNN, RNN, CNN, and LSTM) cannot accurately predict power consumption when it increases or decreases rapidly. Hence, a method to reduce this prediction error is proposed. DNN, RNN, and LSTM were superior when using two-year electricity consumption rather than one-year electricity consumption and weather data.

Keywords: electric load forecasting; deep learning; multilayer perceptron; recurrent neural network; convolution neural network; long short-term memory

1. Introduction

The average global temperature is steadily rising due to global warming. In July 2018, the Korean government announced a policy to reduce 37% greenhouse gases by 2030 [1]. Therefore, the Korean government has proposed technologies and policies to reduce coal power generation and drastically increase renewable energy. In 2018, Korea experienced the 'worst heatwave ever recorded' based on 110 years of recorded meteorological observations. Figure 1 shows the results of the power consumption analysis based on a variety of applications in Korea in 2018 [2]. Figure 1a shows the power consumption that includes industrial (55.70%), general (22.20%), residential (13.90%), agricultural (3.50%), educational (1.60%), late-night (2.40%), and street-lighting applications (0.70%). Figure 1b shows the rate of increase in power consumption by application in 2018 compared to the previous year (2017). The results were analyzed as follows.

1. The rate of power consumption by application increased; this included agricultural (7.30%), residential (6.30%), general (5.10%), educational (4.30%), industrial (2.50%), late-night (2%), and street-lighting applications (0.70%), which are the most sensitive to seasonal factors.
2. The annual domestic power consumption exceeded 70,000 GWh for the first time in 2018. This is expected to increase steadily due to future heat waves.
3. Industrial use increased by 2.5% compared to the previous year, but it accounted for 55.70%. Thus, the power consumption increased considerably to over 2.92 million
GWh. It was determined that the boom in the exports of major energy-consuming industries, such as domestic semiconductors and petrochemicals, was affected.

4. This study proves that seasonal and economic factors influence power consumption [3–5].

![Figure 1](image_url)

**Figure 1.** Power consumption by application in Korea in 2018. (a) The power consumption by application in 2018. (b) The growth rate of the power consumption by application in 2018 compared to 2017.

In addition, the existing electricity demand forecasting studies have been steadily progressing according to the forecast horizon, data characteristics, and data combinations.

1. The forecast horizon of electricity demand is essential because of the maintenance schedule of the energy management system (EMS) [6], the generation of power system utilities, expansion of operations and plans [7], load switching, safety, market demand assessment, cost reduction, and guarantee of continuous power supply [8]. Forecast horizons can be divided into four types [9–11].

   - Very short-term load forecasting (VSTLF) predicts power consumption and demand in real-time by performing predictions for 1 h, 30 min, and 15 min to stably operate power.
   - Short-term load forecasting (STLF) is the most used widely used technique that can perform predictions from 1 h to 1 week, and the forecasting stage is time-based. This prediction plays an essential role in operating the power system and plans the arrival and departure of units and optimal load distribution.
   - Medium-term load forecasting (MTLF) can perform predictions from 1 month to 1 year and includes a daily forecasting step. This prediction is essential for medium-term planning, including the economic operation and repair planning of a system that are directly related to the system’s reliability.
   - Long-term load forecasting (LTLF) performs long-term predictions of over one year. It has a forecast range of over ten years, with a monthly forecasting step. This prediction is essential in network development, such as increasing the number of power plants, transmission lines, and distribution equipment.

2. Data characteristics can be divided into data-driven models (single model) and hybrids (combination model).

   - Data-driven models predict the collected data by using probability and statistics (exponential smoothing [12], autoregressive integrated moving average [13]), classification models (k-nearest neighbors [14], decision trees [15], support vector machines [16], particle swarm optimization [17]), and artificial intelligence (artificial neural networks [18], deep neural networks [19], recurrent neural network [20], long short-term memory [21], and convolution neural networks [22]). This single model has no particular computational complexity,
but its results are unreliable and inaccurate. Hence, the combined model is a good alternative to improve the load prediction accuracy and stability by combining the advantages and disadvantages of a single model [23].

- As for the hybrid model, various models that combine a single model were proposed: (1) EMD–SVM model to predict power demand [24]; (2) random forest-expert model [25], PSO–SVM model [26], and neuro-fuzzy inference–whale optimization algorithm model [27] for short-term load prediction; (3) firefly algorithm–SVR model to predict electrical load related to air conditioning [28]; (4) a combination of CNN and LSTM to extract complex features related to individual energy consumption prediction [29]; and (5) DeepAR, which is an AR-based RNN that combines statistical AR methods and RNN with deep learning [30] (similar to LSTNet [31], TPA [32], DSANet [33], MTNet [34], DSSM [35], DeepTCN [36], DFM-RF [37], DeepGLO [38], and ForecastNet [39]).

3. A model can be categorized as univariate or multivariate according to the data composition (combination). A univariate model performs predictions by only using power consumption. A multivariable model performs predictions by combining the power consumption and weather data (temperature, humidity, sunlight, solar radiation, total cloudiness, wind speed, wind direction, humidity, and vapor pressure). In addition, many recent multivariate models have been used for short-term wind speed prediction. A model applying solar production, temperature, insolation, humidity, and wind speed [40] has been used.

The critical factors for forecasting the building power consumption vary according to the weather and indoor conditions, the size and use of the buildings, and the prediction horizons [41–44]. Among these factors, (1) meteorological conditions are the main predictive factors that can change the indoor conditions and the activities that govern the building power consumption. For example, ambient temperature, cloud cover, humidity, and insolation affect the load pattern due to the comfort of the occupants and adjustment of the lighting levels [45,46]. (2) Commercial buildings that are more than ten times larger than typical residential buildings will most likely consume more electricity than weather variation factors. For example, the load patterns of the buildings vary according to usage and time. This includes single-family houses, multifamily houses, cultural and assembly facilities, religious facilities, education, and research facilities, factories, and warehouse facilities. The varying complexity of load patterns makes it difficult to forecast the building power consumption accurately. (3) Mid/long-term and short-term power consumption forecasting can efficiently operate peer-to-peer (P2P) power transactions and EMS that are suitable for power supply and demand. In particular, changes in consumers’ power consumption patterns are uncertain factors for future power demand. Accurate demand forecasting begins by responding to fluctuations in electricity demand, considering weather/environmental changes, and reflecting recent demand patterns. The response to real-time prediction is brought about by frequency control and economic dispatch. Short-term forecasts account for power generation plans, and mid-term forecasts account for the maintenance of power facilities. Meanwhile, long-term forecasts account for construction plans of generators and transmission networks.

In this study, the seasonal factors and the building power consumption pattern are reflected and proposed to forecast the power consumption accurately.

1. Four models (TM1–TM4) are proposed, and an analysis of the correlation between the seasonal factors, weather data, and power consumption is conducted. These are univariate and multivariate models that have collected power consumption and weather data (temperature, humidity, sunlight, solar radiation, the total cloud cover, wind speed, wind direction, humidity, and vapor pressure) for 1–2 years.

2. The prediction horizon for power consumption forecasting was 24 h, which ensured that power suppliers operated peer-to-peer power transactions and energy management systems efficiently.
3. The proposed four models are compared and analyzed by applying a DNN, RNN, CNN, and LSTM to deep learning, which is currently an issue.

4. Two industrial buildings in the agricultural complex in Naju, Jeollanam-do, were selected to test and verify the proposed method. Building B is an industrial building with constant power consumption, which houses a manufacturing company. Building T is an industrial building with nonuniform power consumption, which houses a livestock processing company. In addition, the experimental data adopted in this study were collected for three years (from 2017 to 2019) by applying industrial electricity real-time rates. The usage was provided by the Korea Electric Power Corporation’s (KEPCO) i-smart [47].

The structure of this study is as follows. Section 2 analyzes the power consumption for Companies B and T’s buildings, which have different power consumption patterns. Section 3 briefly explains the experiments using deep learning methods. Section 4 describes the problems of deep learning, and it explains the proposed method. Section 5 describes the experiment and the analysis of the proposed method. Finally, Section 6 presents the conclusions of this study and the future directions of research.

2. Analysis of the Power Consumption of Companies B and T

Section 2.1 explains the meteorological factors affecting power consumption, which the Korea Meteorological Administration recently announced. Section 2.2 analyzes the power consumption of Companies B and T located in Naju, Jeollanam-do, by season. Finally, Section 2.3 analyzes the correlation between the meteorological data and the power consumption by season.

2.1. Existing Research on the Building Electricity Consumption

In previous studies [48–50], the meteorological factors affecting the building power consumption include the temperature (average, maximum, and minimum), humidity, wind speed, cloudiness, discomfort index, perceived temperature, and precipitation [48]. In Korea, summer (July–August) and winter (December–February) highly correlate with the minimum temperature, cloudiness, a sensible temperature, and power consumption. During the changing seasons (May and October), the correlation between the meteorological factors and the power consumption is relatively low [49]. When forecasting the power consumption using only the temperature, the prediction error was 1.8%. In contrast, when weather factors such as humidity, cloudiness, sensible temperature, wind speed, and precipitation are added, the prediction error is reduced to 1.3%. When forecasting the power consumption, the prediction error can be improved by 25% depending on the combination of meteorological factors; the annual power generation amount can be reduced by approximately 1100 GWh, saving about 120 billion KRW [50]. However, this analysis has the disadvantage in that it may differ slightly from the actual electricity demand forecast since it does not consider variables such as lifestyle, industrial activities, and economic factors.

2.2. Analysis of the Power Consumption by Season

To analyze the power consumption patterns of Companies B and T, the months of April, July, October, and January were selected, which represent Korea’s spring, summer, fall, and winter, respectively. Figure 2 shows a comparison of the power consumption of each month (per 24 h) for Companies B and T during spring (April), summer (July), fall (October), and winter (January). Company B’s building uses electricity from 9 a.m. to 12 p.m. and from 1 to 6 p.m. on weekdays. It does not use electricity during hours other than holidays and working hours. By comparing the power consumption by season, the order is winter > fall > summer > spring. In the winter, the power consumption is high. In contrast, the power consumption is low in the summer (July) since this is the vacation season.
is winter > fall > summer > spring. In the winter, the power consumption is high. In contrast, the power consumption is low in the summer (July) since this is the vacation season. The power consumption of Company T has an irregular pattern. This is because Company T is an agroprocessed food company that consumes electricity according to economic factors (consumer demand) without affecting the season, time, or public holidays.

![Figure 2](https://example.com/figure2.png)

Figure 2. Comparison of the power consumption per day (24 h) between Companies B and T for spring (April), summer (July), fall (October), and winter (January). (a) Company B’s power consumption for spring, summer, fall, and winter; (b) Company T’s power consumption for spring, summer, fall, and winter.

The power consumption of Company T has an irregular pattern. This is because Company T is an agroprocessed food company that consumes electricity according to economic factors (consumer demand) without affecting the season, time, or public holidays.

Table 1 shows the average (average), standard deviation (Std.), and the coefficient of variation (CV) for the seasonal power consumption of Companies B and T. The CV can compare the relative scope; however, it is not enough to calculate the range or variance. Therefore, to compare the power consumption of the different units of measurement, the CV is adopted (Equation (1)):

$$\text{CV} = \frac{\sigma}{\overline{m}}$$

The larger the value of the CV, the larger the relative difference. In addition, $\sigma$ and $\overline{m}$ represent the standard deviation and mean, respectively.

|       | Spring  | Summer  | Fall    | Winter  |
|-------|---------|---------|---------|---------|
| B     | Ave.    | 25.78   | 19.33   | 27.57   | 39.24   |
|       | Std.    | 30.98   | 30.15   | 39.13   | 38.42   |
|       | CV      | 1.16    | 1.26    | 1.09    | 0.97    |
| T     | Ave.    | 124.00  | 114.22  | 122.77  | 119.87  |
|       | Std.    | 31.44   | 24.61   | 26.76   | 26.15   |
|       | CV      | 0.25    | 0.21    | 0.21    | 0.21    |

Company B has a relatively small power consumption than Company T; however, the mean, standard deviation, and CV are large. This is because there is a relatively significant difference between the working hours (9 AM to 6 PM) in which electricity is used and the nonworking hours wherein electricity is not in use. Company T consumes more electricity than Company B. Still, it consumes electricity regardless of the working hours, holidays, and weather factors. Therefore, the standard deviation and CV values are small.
2.3. Correlation between the Seasonal Weather Data and Power Consumption

Figure 3 shows the analysis of the correlation ($R^2$) and trend line ($Y$) of the temperature and humidity by season for the Naju region where Companies B and T are located. The temperature and humidity in the spring, fall, and winter did not seem to have any correlation. The correlation ($R^2$) was determined to be 0.76 for the summer temperature and humidity.

Figures 4 and 5 show the trend line ($Y$) analysis and the correlations ($R^2$) of Companies B and T’s temperature and power consumption by season, respectively; Company T has a higher correlation on average than Company B.

Figure 3. Correlation analysis of the temperature and humidity by season. (a) Spring; (b) Summer; (c) Fall; (d) Winter.

Figure 4. Analysis of the correlation between the temperature and power consumption by season (Company T). (a) Spring; (b) Summer; (c) Fall; (d) Winter.
2.4. Concluding Remarks

Section 2 analyzed the relationships between temperature and humidity, seasonal meteorological factors, and temperature and power consumption. Although weather factors affect power consumption, it was determined that building habits, industrial activities, and economic factors have the greatest influence on power consumption.

3. Deep Learning for Power Consumption

Section 3.1 briefly describes the DNN, LSTM, RNN, and CNN. Section 3.2 presents the comparison and analysis of forecasting the power consumption using deep learning for Companies B and T.

3.1. Deep Learning

3.1.1. Deep Neural Network

DNN is an artificial neural network composed of several hidden layers between the input and output, as shown in Figure 6 [51, 52]. DNNs can model complex nonlinear relationships, such as general artificial neural networks. For example, in a deep neural network structure of an object identification model, each object can be expressed as a hierarchical configuration of the essential elements of an image [53]. Here, the additional layers may aggregate the characteristics of the gradually gathered lower layers. This feature of the deep neural networks enables complex data modeling with fewer units (units, nodes) when compared to similarly performing artificial neural networks [51].

3.1.2. Recurrent Neural Network

RNN refers to a neural network. The connections between the units that constitute an artificial neural network comprise a directed cycle, as shown in Figure 7 [54]. Unlike forward neural networks [20], an RNN can use the memory inside the neural network to process arbitrary inputs. Due to these characteristics, RNNs are used in handwriting recognition fields and show a high recognition rate [55]. A variety of methods are used for structures that can construct RNNs. Hopfield networks [56], Elman networks [57], echo state networks (ESNs) [58], LSTMs [59], bidirectional RNNs [60], continuous-time RNNs (CTRNNs) [61], hierarchical RNNs (HRNNs) [62], and second-order RNNs (SORNNs) [63] are the typical examples. A gradient descent approach, Hessian-free optimization, and
global optimization are typically used to train an RNN. However, the RNN has a scaling issue and is difficult to train when there are many neurons or input units [64,65].

![Figure 6. Structure of DNN.](image)

3.1.3. Concurrent Neural Network

CNN is a multilayer perceptron designed to use minimal preprocessing, as shown in Figure 8 [66]. The CNN comprises one or several convolutional layers. There are general artificial neural network layers on top; it uses weights and pooling layers. Because of this structure, CNN can use the input data of the two-dimensional structure. CNNs perform well in video and audio fields [67,68]. CNNs can also be trained by using standard back-passing methods. CNNs are more accessible to training than the other feed-forward neural network techniques. They have the advantage of using fewer parameters. Recently, a convolutional deep belief network (CDBN) has been developed in deep learning, which is structurally similar to a CNN. It can use the two-dimensional structure of the figure well [69]. Deep belief networks (DBNs) can also take advantage of line training [70]. CDBN provides a general structure that can be used for a variety of image and signal processing techniques. It is used in several benchmark results for standard image data, such as those that belong to the Canadian Institute for Advanced Research (CIFAR) [71].

![Figure 7. Structure of the RNN.](image)

![Figure 8. Structure of the CNN.](image)
3.1.4. Long Short-Term Memory

LSTM is a method used to solve the problem of gradient extinction, which is a disadvantage of the RNN [64,65]. Compared to the basic RNN model, a study has reported that the accuracy is approximately 10% higher when the LSTM model is implemented [72,73]. The interior of the LSTM block comprises a memory cell with a circular structure and three types of gates (input, forget, and output), as shown in Figure 9. Modifications of the LSTM model include a single-layer LSTM [21], multilayer LSTM [74], gated recurrent unit (GRU) [75], bidirectional LSTM [76], encoder–decoder (ED)–LSTM [77], and con–LSTM [78]. Multiple LSTMs have an accuracy higher than single-layer LSTMs.

![Figure 9. Structure of the LSTM.](image)

3.1.5. Persistence Model

The persistence model (PM) is the simplest way of producing a forecast [79,80]. A persistence model assumes that the future value of a time series is calculated under the assumption that nothing changes between the current time and the forecast time (Equation (2)).

\[ \hat{P}_t = O_{t-1} \]  

In Equation (2), \( O_{t-1} \) is the measured power consumption at the time \((t-1)\) and \( \hat{P}_t \) is the predicted value at time \( t \). Its accuracy decreases if the time horizon is longer than 1 h.

3.2. Comparison of the Deep Learning Performance for Companies B and T

In this section, we used the power consumption in the spring, summer, autumn, and winter in 2018 to compare the deep learning performance of Companies B and T. The experimental data were divided into learning data (70%) and test data (30%) and then simulated. Table 2 lists the training and testing options for simulating the PM, DNN, RNN, CNN, and LSTM. In the simulation, the learning rate, loss function (MSE), optimizer (ADAM) [81], and activity function (RELU) [82] were all the same. The remaining options were set for each deep learning method.

| Parameter                      | DNN | RNN | CNN | LSTM |
|--------------------------------|-----|-----|-----|------|
| Number of layers               | 4   | 3   | 6   | 3    |
| Number of neurons              | 100 | 100 | 100 | 100  |
| Number of epochs               | 100 | 100 | 100 | 100  |
| Number of convolution filters  | -   | -   | 64  | -    |
| Convolution of kernel size     | -   | -   | 1   | -    |
| Learning rate                  | 0.05| 0.05| 0.05| 0.05 |
| Loss function                  | MSE | MSE | MSE | MSE  |
| Optimizer                      | ADAM| ADAM| ADAM| ADAM |
| Activation function            | RELU| RELU| RELU| RELU |
Table 3 compares the persistence model and the deep learning performance of Companies B and T by season. Figure 10 shows comparison of the persistence model and the deep learning performance of Companies B and T in April 2018. In addition, the average (Ave.), standard deviation (Std.), CV, root-mean-square error (RMSE) \[83\], and mean absolute percentage error (MAPE) \[84\] were used for the prediction errors. The values with excellent performance for each deep learning method are shown in bold. The RMSE and MAPE are shown in Equations (3) and (4), respectively.

\[
RMSE = \sqrt{\frac{1}{n} \sum (O_i - P_i)^2}
\]

\[
MAPE = \frac{100}{n} \sum \left| \frac{O_i - P_i}{O_i} \right|
\]

In Equations (3) and (4), \(n\) is the total number that is to be predicted, \(O_i\) is the actual value, and \(P_i\) is the predicted value. For RMSE and MAPE, the smaller the number, the better the performance.

Table 3. Average, standard deviation, and coefficient of variation for the power consumption of Companies B and T.

| Season | Class | Metrics | B        | T        |        |        |        |        |        |        |
|--------|-------|---------|----------|----------|--------|--------|--------|--------|--------|--------|
|        |       |         |         |          | PM     | DNN    | RNN    | CNN    | LSTM   | PM     | DNN    | RNN    | CNN    | LSTM   |
|        | Ave.  |         | 31.76   | 29.46   | 28.64  | 30.00  | 29.17  | 139.41 | 137.23 | 138.76 | 137.13 | 131.76 |
|        | Std.  |         | 41.59   | 35.14   | 34.52  | 34.22  | 34.74  | 27.97  | 26.33  | 26.42  | 26.31  | 22.46  |
|        | CV    |         | 1.31    | 1.19    | 1.21   | 1.14   | 1.19   | 0.2    | 0.20   | 0.19   | 0.19   | 0.17   |
|        | RMSE  |         | 20.65   | 20.75   | 20.03  | 23.08  | 20.66  | 19.34  | 19.27  | 19.08  | 20.62  | 20.03  |
|        | MAPE  |         | 43.07   | 56.61   | 57.91  | 98.12  | 53.95  | 10.81  | 10.68  | 10.76  | 11.07  | 10.60  |
| Spring | Ave.  |         | 22.12   | 29.46   | 28.64  | 30.00  | 29.17  | 114.5  | 113.94 | 107.71 | 116.58 | 112.58 |
|        | Std.  |         | 30.99   | 35.14   | 34.52  | 34.22  | 34.74  | 22.65  | 19.81  | 19.50  | 18.78  | 15.90  |
|        | CV    |         | 1.4     | 1.19    | 1.21   | 1.14   | 1.19   | 0.2    | 0.17   | 0.18   | 0.16   | 0.14   |
|        | RMSE  |         | 15.09   | 20.75   | 20.03  | 23.08  | 20.66  | 21.43  | 18.98  | 19.15  | 22.81  | 17.88  |
|        | MAPE  |         | 36.59   | 56.61   | 57.91  | 98.12  | 53.95  | 13.27  | 12.18  | 12.65  | 15.76  | 11.56  |
| Summer | Ave.  |         | 29.13   | 29.20   | 28.74  | 28.36  | 32.22  | 127.87 | 126.32 | 129.51 | 124.44 | 124.31 |
|        | Std.  |         | 37.76   | 34.35   | 33.80  | 31.36  | 33.56  | 21.36  | 18.69  | 19.19  | 17.11  | 16.43  |
|        | CV    |         | 1.3     | 1.18    | 1.18   | 1.11   | 1.04   | 0.17   | 0.15   | 0.15   | 0.14   | 0.13   |
|        | RMSE  |         | 18.70   | 17.82   | 17.72  | 19.86  | 15.93  | 18.47  | 17.73  | 17.82  | 19.65  | 17.16  |
|        | MAPE  |         | 42.97   | 62.50   | 62.58  | 83.04  | 85.82  | 10.74  | 10.11  | 10.62  | 11.47  | 9.62   |
| Fall   | Ave.  |         | 41.36   | 39.42   | 35.91  | 38.86  | 40.72  | 121.45 | 124.58 | 125.99 | 122.35 | 115.87 |
|        | Std.  |         | 37.58   | 33.05   | 29.98  | 31.39  | 34.02  | 25.76  | 23.17  | 23.8   | 21.86  | 15.63  |
|        | CV    |         | 0.91    | 0.84    | 0.83   | 0.81   | 0.84   | 0.21   | 0.19   | 0.19   | 0.18   | 0.13   |
|        | RMSE  |         | 17.27   | 16.24   | 16.91  | 19.05  | 15.93  | 18.79  | 20.12  | 20.30  | 21.89  | 19.26  |
|        | MAPE  |         | 22.88   | 25.00   | 22.93  | 36.81  | 28.77  | 12.26  | 13.84  | 14.12  | 14.71  | 11.52  |
Table 3 compares the persistence model and the deep learning performance of Companies B and T by season. Figure 10 shows comparison of the persistence model and the deep learning performance of Companies B and T in April 2018. In addition, the average (Ave.), standard deviation (Std.), CV, root-mean-square error (RMSE) [83], and mean absolute percentage error (MAPE) [84] were used for the prediction errors. The values with excellent performance for each deep learning method are shown in bold. The RMSE and MAPE are shown in Equations (3) and (4), respectively.

\[
RMSE = \sqrt{\frac{1}{n} \sum (O_j - P_j)^2} \\
\text{MAPE} = \frac{100}{n} \sum \frac{|O_j - P_j|}{O_j}
\]

In Equations (3) and (4), \(n\) is the total number that is to be predicted, \(O_j\) is the actual value, and \(P_j\) is the predicted value. For RMSE and MAPE, the smaller the number, the better the performance.

(a) Comparison of the persistence model and the deep learning performance (Company B); (b) Comparison of the persistence model and the deep learning performance (Company T).

Table 4 shows the average RMSE and MAPE values for the four seasons when using persistence model and deep learning. The RMSE of Company B with large CV fluctuations is: LSTM (16.05) > RNN (17.04) > DNN (17.11) > PM (17.93) > CNN (19.74). However, for MAPE, it is: PM (36.38) > RNN (54.97) > DNN (56.18) > LSTM (63.94) > CNN (93.74). The RMSE of Company T with small CV fluctuations is: LSTM (18.58) > DNN (19.02) > RNN (19.08) > PM (19.51) > CNN (21.24). For MAPE, LSTM (10.08) > DNN (11.70) > PM (11.77) > RNN (12.03) > CNN (13.25). Thus, Company B with large CV fluctuations has different priorities depending on the prediction error method, but Company T with small CV fluctuations has an excellent LSTM regardless of the prediction error method.

Table 4. Performance comparison of the average RMSE and MAPE values for four seasons by deep learning for Companies B and T.

| Metrics | Class | Company B | Company T |
|---------|-------|-----------|-----------|
|         | PM    | DNN       | RNN       | CNN       | LSTM     | PM    | DNN       | RNN       | CNN       | LSTM     |
| RMSE    | 17.93 | 17.11     | 17.04     | 19.74     | 16.05    | 19.51  | 19.02     | 19.08     | 21.24     | 18.58    |
| MAPE    | 36.38 | 56.18     | 54.97     | 93.74     | 63.94    | 11.77  | 11.70     | 12.03     | 13.25     | 10.82    |

4. Proposed Method

Section 4.1 analyzes the power consumption patterns of Companies B and T, and it describes the problems. Section 4.2 proposes the prediction error correction to solve the problems when simulating traditional deep learning for Companies B and T.
4.1. Problem of Existing Deep Learning

Section 2 analyzes the relationship between the weather data and power consumption by season. This analysis reveals that the weather data, lifestyle, industrial activity, and economic factors affect power consumption. For example, Figure 11 compares the electricity usage pattern analysis for the daily averages by season for Companies B and T. Figure 11a shows that Company B has a constant power consumption pattern. Still, it fluctuates rapidly, depending on the period of power usage and lifestyle. Thus, the power consumption increases during the intensive working hours from 8 AM to 6 PM. Still, the power consumption decreases sharply during the rest of the hours (7 PM to 8 AM) and lunchtime (12 PM to 1 PM). Unlike Company B, as shown in Figure 11b, Company T does not have a uniform power consumption pattern. Still, the power consumption does not significantly increase or decrease.

Figure 11. Comparison of the power consumption patterns for the daily average by the season for Companies B and T. (a) Company B; (b) Company T.

4.2. Proposed Forecasting Error Correction

The application of DNN, RNN, CNN, and LSTM in different seasons revealed that the performance of the LSTM was excellent for deep learning. However, as shown in Figure 11a, when the power consumption increases and decreases rapidly, LSTM cannot accurately predict the power consumption; therefore, this study proposes a method to correct the prediction error when the power consumption rapidly increases or decreases. Figure 12 shows the flowchart proposed in this study.

First, as a data collection and preprocessing process, the electricity and weather data of Companies B and T in Naju were collected over 1 h intervals and preprocessed for simulation. Second, after splitting the electricity and weather data into four test models, each was divided into training data (70%) and test data (30%). Third, the training data of four test models were simulated using deep learning (DNN, RNN, CNN, and LSTM). The proposed prediction error correction is applied to the training data of four test models and simulated. Fourth, the learning deep learning module was applied to the test data using four test models. Finally, the performances of the traditional and proposed methods are compared with the prediction error measurement (RMSE, MAPE).
4.2. Proposed Forecasting Error Correction

The application of DNN, RNN, CNN, and LSTM in different seasons revealed that the performance of the LSTM was excellent for deep learning. However, as shown in Figure 11a, when the power consumption increases and decreases rapidly, LSTM cannot accurately predict the power consumption; therefore, this study proposes a method to correct the prediction error when the power consumption rapidly increases or decreases. Figure 12 shows the flowchart proposed in this study.

4.2.1. Data Collection and Preprocessing

The electricity and weather data were collected from 2017 to 2018 for Companies B and T in Naju, Jeollanam-do. The electricity data were collected from Companies B and T in one-hour intervals using i-Smart, which provides a real-time power consumption service from KEPCO. The weather data were collected in one-hour intervals from the Naju Regional Meteorological Administration. The collected electricity and weather data were preprocessed monthly (January, April, August, and October).

4.2.2. Proposed Four Models

This section proposes four models (TM1–TM4), as shown in Table 5, which combine the power consumption and weather data that affect the power consumption prediction. For TM1, one variable was univariate; for TM2 to TM4, multivariate variables with two or more variables were adopted. TM1 was the most frequently used power consumption variable for estimating the existing power consumption. The power consumption data for Companies B and T were collected by season in one-hour units in 2018. TM2 predicts the power consumption by collecting data over two years (from 2017 to 2018). TM3 comprises the 2018 power consumption and the weather data (e.g., temperature, humidity, sunlight, and cloud cover). Finally, TM4 includes the two-year power consumption and the weather data (e.g., temperature, humidity, sunshine, cloud cover, precipitation, wind direction, wind speed, and vapor pressure).
Table 5. Proposed four models.

| Variables | Model Name | Model Description |
|-----------|------------|-------------------|
| Univariate | TM1        | Power consumption in 2018 |
|           | TM2        | Power consumption for two years (2017, 2018) |
| Multivariate | TM3     | Power consumption in 2018, temperature (°C), humidity (%), solar radiation (MJ/m²), cloud cover |
|           | TM4        | Power consumption for two years (2017, 2018), temperature (°C), humidity (%), solar radiation (MJ/m²), cloud cover, precipitation, wind speed, wind direction, vapor pressure |

Figure 13 shows that the four test models split the training and test data ratio to 70:30. Four test models separated by the training data were simulated for each deep learning method using Table 2. The training data had different input variables for each TM, but all have synchronized data in units of 1 h. The test data of the four TMs predict the power consumption from 22 April 2018, 1:00 (a.m.), to 31 April 2018, 12:00 (p.m.). Figure 13a is a simulation of the April 2018 (spring) power data, with learning data from 1 April 2018, 1:00 (a.m.), to 21 April 2018, 12:00 (p.m.). Figure 13b uses data from April 2017 and April 2018, wherein two variables were synchronized each hour. The training data used in TM2 are from 1 April 2017, 1:00 (a.m.), and 1 April 2018, 1:00 (a.m.), to 21 April 2018, 12:00 (p.m.). Figure 13c,d shows five (power consumption, temperature, humidity, solar radiation, and cloud cover of April 2018) and nine variables (power consumption of two years, temperature, humidity, solar radiation, cloud cover, precipitation, wind speed, wind direction, and vapor pressure of April 2018), respectively.

![Figure 13](image)

**Figure 13.** Data partition for training and testing by four TMs. (a) TM1; (b) TM2; (c) TM3; (d) TM4.

4.2.3. Proposed Error Correction

Section 4.1 states that if the power consumption increases or decreases rapidly, it cannot be accurately predicted. Therefore, we propose a prediction error correction method shown in Equation (5) by considering the difference between the previous and current power consumptions.

\[
\hat{P}_i = O_i - P_i + \left| \frac{O_i - O_{i-1}}{O_i} \right|
\]  

(5)

In Equation (5), \(O_i\) and \(O_{i-1}\) represent the actual power consumption for \(i\) and \(i - 1\) h, respectively. \(P_i\) is the predicted power consumption, and \(\hat{P}_i\) is the prediction error correction proposed in this study. Finally, the learned deep learning modules simulated four test data points for each deep learning method.
4.2.4. Prediction Error Measurement

This study adopted RMSE and MAPE to measure the prediction error of the existing and proposed methods using Equations (3) and (4) in Section 3.2. RMSE is the most commonly used prediction error measurement method [83]. MAPE is the most common measurement used to forecast error and works best if there are no extremes in the data (and no zeros) [84].

5. Simulation Results and Analysis

5.1. Test Environment

To verify the traditional and proposed methods, the experiments were performed on a workstation computer equipped with an Intel Xeon (R) W-2133, a 3.60 GHz CPU, and 32 GB of RAM (Dell Precision 5820 Tower Workstation). The operating system was Windows 10 Pro for the workstations (64 bit). The traditional and proposed methods were implemented using deep learning libraries provided by Tensor flow [85] and Keras [86].

5.2. Company B with a Constant Power Consumption Pattern

Section 2.2 states that Company B has a constant power consumption pattern. Table 6 shows the results of applying traditional deep learning and the proposed method to four TMs. Although there is a slight difference in the deep learning performance based on the month, CNN has the worst performance among the four TMs. However, the DNN and RNN exhibited an excellent TM performance. LSTM showed that the TM1, TM2, and TM3 performances are similar to those of DNN and RNN; however, TM4 did not perform well. Therefore, the power consumption prediction will perform well if the test model adopts TM1 or TM2 and simulates DNN, RNN, and LSTM.

Figure 14 shows a comparison of MAPE (%) for the traditional and the proposed methods for Company B in April 2018. Figure 14a is a performance comparison of four TMs using deep learning. The error between the traditional and the proposed methods appeared in TM4, TM3, TM2, and TM1. For DNN, RNN, and LSTM, the more variables were adopted, the greater the error. Figure 14b is a comparison of deep learning performance by TM. The error between the traditional and the proposed method appeared in the order CNN, LSTM, RNN, and DNN. In particular, CNN performed poorly in all TMs. For DNN, RNN, and LSTM, the smaller the variable, the smaller the error.

Figure 15 shows the results of comparing the performance of MAPE (%) by TMs and deep learning for the proposed method for Company B in January 2018. Figure 15a shows the deep learning performance of TMs, which exhibited the following order: DNN, RNN, LSTM, and CNN. Figure 15b shows the results of four test models for each deep learning method; the performance was excellent in the order of TM1, TM2, TM3, and TM4.

Figure 16 shows the performance analysis of the proposed DNN, RNN, CNN, and LSTM for the four TMs in January 2018. Figure 16a shows the proposed DNN performance analysis for the four TMs. The predicted values (DNN-TM1 and DNN-TM2), in which TM1 and TM2 were applied, have predictions similar to the actual value (power). The predicted values (DNN-TM3 and DNN-TM4) when applying TM3 and TM4 from 97 to 121 times showed that the predicted values were not similar to the actual values. Figure 16b shows the proposed RNN performance analysis for the four TMs. Like the proposed DNN results, the predicted values of the proposed RNNs (RNN-TM1 and RNN-TM2), in which TM1 and TM2 were applied, have been predicted to be similar to the actual values. However, the predicted value of TM3 (RNN-TM3) exhibits a higher error (RMSE: 25.70) from the actual value than that shown by the predicted value of TM4 (RNN-TM4); thus, the performance was poor. Figure 16c shows the performance analysis of the proposed CNN for the four TMs. Similar to the proposed DNN results, the predicted values (CNN-TM1, CNN-TM2) of the CNNs, in which TM1 and TM2 were applied, have been predicted to be similar to the actual values. However, the predicted value of TM4 (CNN-TM4) was too high (RMSE: 19.44) from the actual value; thus, the performance was not good compared to that of TM3 (CNN-TM3). Figure 16d shows the proposed LSTM performance analysis for the four TMs.
Similar to the proposed DNN result, the predicted value (LSTM-TM1) of the LSTM, in which TM1 was applied, has been predicted to be like the actual value. However, the TM3 predicted value (LSTM-TM3) exhibited a high error (RMSE: 25.92) because the difference from the actual value was too large.

Figure 17 illustrates a performance comparison of DNN, RNN, CNN, and LSTM for TM2, which has an excellent performance among the four TMs. The deep learning performance comparison results for TM2 are in the following order: DNN-TM2 (RMSE: 15.05) > RNN-TM2 (RMSE: 15.46) > LSTM-TM2 (RMSE: 17.36) > CNN-TM2 (RMSE: 18.23).

Table 6. Comparison of the performance of the traditional and the proposed method by deep learning (Company B).

| Mon. | TM | Metrics | Traditional Method | Proposed Method |
|------|----|---------|--------------------|-----------------|
|      |    |         | DNN | RNN | CNN | LSTM | DNN | RNN | CNN | LSTM |
| Jan. | TM1 | RMSE    | 16.24 | 16.91 | 19.05 | 15.93 | 15.20 | 15.86 | 18.12 | 14.92 |
|      |     | MAPE    | 25.00 | 22.93 | 36.81 | 28.77 | 24.75 | 22.70 | 36.44 | 28.49 |
| TM2  |     | RMSE    | 16.10 | 16.48 | 19.25 | 18.27 | 15.05 | 15.46 | 18.23 | 17.36 |
|      |     | MAPE    | 26.24 | 27.57 | 40.27 | 33.35 | 25.98 | 27.30 | 39.87 | 33.01 |
| TM3  |     | RMSE    | 16.48 | 25.70 | 18.58 | 25.92 | 15.43 | 24.78 | 17.53 | 25.21 |
|      |     | MAPE    | 36.77 | 39.73 | 39.32 | 31.46 | 36.40 | 39.33 | 38.93 | 31.14 |
| TM4  |     | RMSE    | 16.78 | 19.79 | 19.44 | 19.48 | 15.73 | 19.19 | 18.46 | 18.47 |
|      |     | MAPE    | 39.82 | 39.25 | 44.04 | 36.00 | 39.42 | 38.86 | 43.60 | 35.64 |
| Apr. | TM1 | RMSE    | 20.75 | 20.03 | 23.09 | 20.66 | 18.81 | 18.87 | 22.30 | 19.48 |
|      |     | MAPE    | 56.61 | 57.91 | 98.12 | 53.95 | 56.05 | 57.34 | 97.14 | 53.42 |
| TM2  |     | RMSE    | 23.02 | 21.00 | 23.54 | 22.54 | 18.82 | 19.78 | 22.63 | 21.33 |
|      |     | MAPE    | 63.74 | 61.81 | 92.32 | 56.95 | 63.11 | 61.20 | 91.40 | 56.38 |
| TM3  |     | RMSE    | 28.36 | 29.47 | 26.54 | 30.35 | 18.02 | 28.42 | 25.35 | 32.47 |
|      |     | MAPE    | 80.18 | 68.14 | 88.10 | 71.44 | 79.38 | 67.46 | 87.22 | 70.72 |
| TM4  |     | RMSE    | 21.65 | 21.70 | 25.86 | 26.31 | 18.52 | 20.81 | 23.93 | 25.64 |
|      |     | MAPE    | 87.45 | 89.23 | 102.13 | 110.17 | 86.58 | 88.34 | 101.11 | 109.07 |
| Jul. | TM1 | RMSE    | 13.66 | 13.52 | 17.00 | 11.71 | 12.49 | 12.38 | 16.19 | 11.12 |
|      |     | MAPE    | 80.64 | 76.48 | 157.02 | 87.23 | 79.84 | 75.72 | 155.45 | 86.36 |
| TM2  |     | RMSE    | 13.47 | 14.28 | 16.88 | 13.17 | 12.30 | 13.44 | 15.80 | 12.47 |
|      |     | MAPE    | 91.48 | 88.76 | 162.92 | 69.06 | 90.57 | 87.87 | 161.29 | 68.37 |
| TM3  |     | RMSE    | 13.15 | 14.37 | 13.45 | 14.83 | 14.58 | 14.00 | 13.44 | 22.63 |
|      |     | MAPE    | 100.59 | 63.42 | 131.87 | 60.92 | 99.58 | 62.78 | 130.55 | 60.31 |
| TM4  |     | RMSE    | 13.22 | 15.94 | 15.29 | 18.28 | 14.24 | 57.99 | 14.56 | 17.61 |
|      |     | MAPE    | 126.81 | 138.84 | 155.12 | 115.23 | 125.54 | 137.45 | 153.57 | 114.08 |
| Oct. | TM1 | RMSE    | 17.82 | 17.72 | 19.86 | 15.93 | 16.81 | 16.74 | 18.78 | 15.18 |
|      |     | MAPE    | 62.50 | 62.58 | 83.04 | 85.82 | 61.88 | 61.96 | 82.21 | 84.97 |
| TM2  |     | RMSE    | 17.72 | 17.74 | 19.70 | 14.49 | 16.72 | 16.74 | 18.76 | 15.71 |
|      |     | MAPE    | 62.25 | 62.17 | 94.10 | 64.75 | 61.63 | 61.55 | 93.16 | 64.10 |
| TM3  |     | RMSE    | 16.99 | 17.48 | 18.40 | 17.03 | 16.16 | 16.63 | 17.41 | 16.47 |
|      |     | MAPE    | 78.87 | 84.95 | 104.70 | 95.08 | 78.09 | 84.10 | 103.65 | 94.13 |
| TM4  |     | RMSE    | 17.43 | 20.68 | 19.67 | 24.53 | 17.02 | 20.31 | 18.85 | 24.30 |
|      |     | MAPE    | 86.53 | 91.25 | 94.88 | 116.60 | 85.66 | 90.33 | 93.93 | 115.43 |
Figure 14. Performance comparison of MAPE (%) by traditional and proposed methods for Company B (April 2018). (a) Comparison of four TMs by deep learning; (b) Deep learning comparison by four TMs.

Figure 15. Performance comparison of MAPE (%) by four test models and deep learning for Company B (January 2018). (a) Deep learning comparison by four TMs. (b) Comparison of four TMs by deep learning.

5.3. Company T with the Irregular Power Consumption Pattern

Table 7 shows the profound learning performance results for Company T with irregular power consumption patterns. By comparing the performance of the four test models, the prediction error for each test model was not significant for Company T, unlike Company B. In other words, Company T has a slight difference in the deep learning performance based on the month, but similar results are observed for Company B, as follows: (1) The DNN and RNN are the four TMs with excellent performance. (2) The performances of the LSTM and CNN are similar to DNN and RNN in TM1–TM3, but TM4 did not perform well. Therefore, although Company T has an irregular power consumption pattern, it has an excellent performance when simulating DNN, RNN, and LSTM by adopting TM1 or TM2 for predicting the power consumption.

Figure 18 shows a comparison of MAPE (%) for the traditional and the proposed methods for Company T in April 2018. Figure 18a is a performance comparison of four TMs using deep learning. The error between the traditional and the proposed methods was almost similar to that observed for the TM performance, unlike Company B. Similar to Company B, Company T’s CNN performed poorly in all TMs. Figure 18b shows a comparison of the deep learning performance exhibited by TM. The error between the traditional and the proposed methods appeared in the following descending order: CNN, LSTM, RNN, and DNN. DNN, RNN, and LSTM performed similarly regardless of the number of variables.
value. However, the TM3 predicted value (LSTM-TM3) exhibited a high error (RMSE: 25.92) because the difference from the actual value was too large.

Figure 16. Performance comparison of the four test models when applying deep learning (January 2018). (a) DNN performance comparison for the four TMs. (b) RNN performance comparison for the four TMs. (c) CNN performance comparison for the four TMs. (d) LSTM performance comparison for the four TMs.

Figure 17 illustrates a performance comparison of DNN, RNN, CNN, and LSTM for TM2, which has an excellent performance among the four TMs. The deep learning performance comparison results for TM2 are in the following order: DNN-TM2 (RMSE: 15.05) > RNN-TM2 (RMSE: 15.46) > LSTM-TM2 (RMSE: 17.36) > CNN-TM2 (RMSE: 18.23).

Figure 19 shows the results of comparing the MAPE (%) performance by using TMs and deep learning for the method proposed for Company T in January 2018. Figure 19a shows the deep learning performance of TMs, which exhibited the following order: DNN, RNN, LSTM, and CNN. Figure 19b shows the results of four test models for each deep learning model; the performance exhibited the following order: TM2, TM3, TM1, and TM4.

Figure 20 shows the DNN, RNN, CNN, and LSTM proposed for the four TMs in January 2018 (winter) for Company T. Figure 20a shows the proposed DNN performance analysis for the four TMs. The four TMs displayed excellent performance. Thus, a building such as Company T, whose power consumption pattern is not constant, will perform well
Table 7. Comparison of the performance of the traditional and the proposed method by deep learning (Company T).

| Mon. | TM       | Metrics | Traditional Method | Proposed Method |
|------|----------|---------|--------------------|-----------------|
|      |          |         | DNN               | RNN             | CNN | LSTM | DNN | RNN | CNN | LSTM |
| Jan. | TM1      | RMSE    | 20.12             | 20.30           | 21.89| 19.26| 19.12| 19.30| 20.83| 18.25|
|      |          | MAPE    | 13.84             | 14.12           | 14.71| 11.52| 13.71| 13.98| 14.50| 11.41|
| TM2  | RMSE     | 19.04   | 19.29             | 20.46           | 20.69| 18.03| 18.29| 19.47| 19.70|
|      | MAPE     | 11.74   | 12.80             | 13.65           | 13.09| 11.62| 12.67| 13.51| 12.97|
| TM3  | RMSE     | 21.23   | 20.6              | 20.22           | 22.78| 20.23| 19.37| 19.36| 21.82|
|      | MAPE     | 15.46   | 12.62             | 13.37           | 13.85| 15.31| 12.50| 13.29| 13.74|
| TM4  | RMSE     | 19.19   | 25.71             | 22.51           | 26.13| 18.20| 24.73| 21.64| 25.20|
|      | MAPE     | 12.38   | 14.87             | 14.67           | 16.82| 12.26| 14.75| 14.59| 16.70|
| Apr. | TM1      | RMSE    | 19.27             | 19.08           | 20.62| 20.03| 18.26| 18.07| 19.57| 19.02|
|      |          | MAPE    | 10.68             | 10.76           | 11.09| 10.60| 10.57| 10.65| 10.94| 10.50|
| TM2  | RMSE     | 18.72   | 19.26             | 20.78           | 19.51| 17.71| 18.24| 19.72| 18.50|
|      | MAPE     | 10.26   | 10.43             | 10.91           | 10.30| 10.16| 10.33| 10.78| 10.19|
| TM3  | RMSE     | 19.31   | 19.97             | 18.64           | 19.77| 18.30| 18.97| 17.58| 18.77|
|      | MAPE     | 10.08   | 10.62             | 10.34           | 10.63| 9.98 | 10.52| 10.20| 10.52|
| TM4  | RMSE     | 19.00   | 19.96             | 19.11           | 24.20| 17.99| 18.96| 18.05| 23.23|
|      | MAPE     | 10.48   | 11.25             | 10.74           | 13.38| 10.37| 11.14| 10.61| 13.27|
| Jul. | TM1      | RMSE    | 18.98             | 19.15           | 22.81| 17.88| 17.98| 18.15| 21.78| 16.90|
|      |          | MAPE    | 12.18             | 12.65           | 15.76| 11.57| 12.06| 12.54| 15.55| 11.45|
| TM2  | RMSE     | 18.38   | 18.85             | 21.54           | 17.56| 17.37| 17.84| 20.49| 16.58|
|      | MAPE     | 12.12   | 12.72             | 14.48           | 12.27| 17.37| 17.84| 20.49| 16.58|
| TM3  | RMSE     | 16.99   | 16.13             | 20.27           | 17.15| 15.98| 15.11| 19.22| 16.14|
|      | MAPE     | 11.53   | 11.08             | 13.93           | 11.42| 11.41| 10.96| 13.74| 11.32|
| TM4  | RMSE     | 17.97   | 19.43             | 21.63           | 22.29| 16.97| 18.44| 20.60| 21.34|
|      | MAPE     | 13.06   | 13.31             | 15.70           | 15.37| 12.92| 13.20| 15.50| 15.25|
Comparison of four TMs by deep learning; Figure 18. Performance comparison of MAPE (%) by traditional and proposed methods for Company T (April 2018). (a) Comparison of four TMs by deep learning; (b) Deep learning comparison by four TMs.

Figure 19. Performance comparison of MAPE (%) by four test models and deep learning for Company T (January 2018). (a) Deep learning comparison by four TMs. (b) Comparison of four TMs by deep learning.

Figure 21 compares the performances of the proposed DNN, RNN, CNN, and LSTM for TM2, which exhibits an excellent performance among the four TMs. Based on the deep learning performance comparison for TM2, it was shown in the order of DNN-TM2 (RMSE: 18.03) > RNN-TM2 (RMSE: 18.29) > CNN-TM2 (RMSE: 19.47) > LSTM-TM2 (RMSE: 19.70).
The four TMs displayed excellent performance. Thus, a building such as Company T, whose power consumption pattern is not constant, will perform well if the power consumption is predicted using the existing power consumption and weather data. In Figure 20b–d, the predicted values of RNN, CNN, and LSTM are applied with TM1–TM3, which have a similar performance to the actual values. However, the predicted values of RNN, CNN, and LSTM applied with TM4 have a large error difference when compared to the measured values; thus, the performance was poor.

Figure 20. Performance comparison of the four test models by applying deep learning (January 2018). (a) DNN performance comparison for the four TMs. (b) RNN performance comparison for the four TMs. (c) CNN performance comparison for the four TMs.
6. Conclusions and Future Research

Forecasting the building power consumption is necessary to improve the energy efficiency of buildings and overcome the power peak problem or energy crisis that occur every year worldwide. Building owners and emergency medical services operators need to manage their buildings’ electrical energy consumption. Since electrical energy is the main form of energy consumed in commercial buildings, the ability to predict the electrical energy consumption would benefit building owners and operators. Data-driven models for energy forecasting have been extensively studied over the past few decades due to their improved performance, robustness, and ease of deployment. Artificial neural networks are one of the most popular data-driven approaches applied to date among the different models.

In this study, power demand prediction was proposed by using deep learning for two industrial buildings with different power consumption patterns. At this time, four TMs were constructed and simulated while considering the influences on power consumption. This includes the industrial activities, economic factors, resident living patterns, and weather data for buildings, which represent the primary power consumption.

Although the performances are different for each of the four TMs and deep learning methods, TM1 (power consumption for one year) and TM2 (power consumption for two years), which do not use weather data, show excellent performance when using DNN, RNN, and LSTM. In contrast, the performance was poor if weather data that did not affect power consumption were used (such as TM4). Therefore, it is more critical to select weather data (temperature, humidity, solar radiation, and cloud cover) that only affects power consumption, such as TM3.

In future research, the proposed method in this article will be applied worldwide by overcoming the limitations of the Korean region.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the National Program for Excellence in SW, supervised by the IITP (Institute for Information & Communications' Technology Planning & Evaluation) in 2021 (2021-D-01409). I would like to thank Wonhee Cho, Hyunjeong Yang, and Evan Kim, who collaborate this article.

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

1. 2030 National Greenhouse Gas Reduction Target; Ministry of Industry: Sejong City, Korea, 2020. Available online: https://english.motie.go.kr/en/am/introduction/introduction.jsp (accessed on 27 October 2021).

2. KOSIS: Korean Statistical Information Service. Available online: https://kosis.kr/statHtml/statHtml.do?orgId=771&tblId=DT_77101_G001002 (accessed on 12 March 2021).

3. Kandananond, K. Forecasting electricity demand in Thailand with an artificial neural network approach. *Energies* **2011**, *4*, 1246–1257. [CrossRef]

4. Brown, R.E.; Kooi, J.G. Electricity use in California: Past trends and present usage patterns. *Energy Policy* **2003**, *31*, 849–864. [CrossRef]

5. Niu, S.; Jia, Y.; Wang, W.; He, R.; Hu, L.; Liu, Y. Electricity consumption and human development level: A comparative analysis based on panel data for 50 countries. *Int. J. Electr. Power Energy Syst.* **2013**, *53*, 338–347. [CrossRef]

6. Sideratos, G.; Ikonomopoulos, A.; Hatzigiouyiou, N.D. A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks. *Electr. Power Syst. Res.* **2020**, *178*, 106025. [CrossRef]

7. Maldonado, S.; González, A.; Crone, S. Automatic time series analysis for electric load forecasting via support vector regression. *Appl. Soft Comput.* **2019**, *83*, 105616. [CrossRef]

8. Yang, A.; Li, W.; Yang, X. Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines. *Knowl.-Based Syst.* **2019**, *163*, 159–173. [CrossRef]

9. Raza, M.Q.; Khosravi, A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renew Sustain Energy Rev.* **2015**, *50*, 1352–1372. [CrossRef]

10. Willis, H.L.; Northcote-Green, J.E.D. Spatial electric load forecasting: A tutorial review. *Proc. IEEE* **1983**, *71*, 232–253. [CrossRef]

11. Enea, M. A review of machine learning algorithms used for load forecasting at micro-grid level. In *Sinteza 2019—International Scientific Conference on Information Technology and Data Related Research*; Singidunun University: Belgrade, Serbia, 2019; pp. 452–458.

12. Brown, R.G. Smoothing Forecasting and Prediction of Discrete Time Series; Prentice-Hall: Englewood Cliffs, NJ, USA, 1963.

13. Ohtsuka, Y.; Oga, T.; Kakamou, K. Forecasting electricity demand in Japan: A Bayesian spatial autoregressive ARMA approach. *Comput. Stat. Data Anal.* **2010**, *54*, 2721–2735. [CrossRef]

14. Fix, E.; Hodges, J.L., Jr. *Discriminatory Analysis-Nonparametric Discrimination: Consistency Properties*; International Statistical Institute: Voorburg, The Netherlands, 1989; Volume 57, pp. 238–247.

15. Yu, Z.; Haghighat, F.; Fung, B.C.M.; Yoshino, H. A decision tree method for building energy demand modeling. *Energy Build.* **2010**, *42*, 1637–1646. [CrossRef]

16. Dong, B.; Cao, C.; Lee, S.E. Applying support vector machines to predict building energy consumption in tropical region. *Energy Build.* **2005**, *37*, 545–553. [CrossRef]

17. Bonyadi, M.R.; Michalewicz, Z. Particle swarm optimization for single objective continuous space problems: A review. *Evol. Comput.* **2017**, *25*, 1–54. [CrossRef]

18. Kalogirou, S.A.; Neocleous, C.C.; Schizas, C.N. Building heating load estimation using artificial neural networks. In Proceedings of the 17th International Conference on Parallel Architectures and Compilation Techniques, San Francisco, CA, USA, 10–14 November 1997.

19. Bagnasco, A.; Fresi, F.; Saviozzi, M.; Silvestro, F.; Vinci, A. Electrical consumption forecasting in hospital facilities: An application case. *Energy Build.* **2015**, *103*, 261–270. [CrossRef]

20. Graves, A.; Liwicki, M.; Fernandez, F.; Bertolami, R.; Bunke, H.; Schmidhuber, J. A Novel Connectionist System for Improved Unconstrained Handwriting Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 855–868. [CrossRef]

21. Gers, F.; Schmidhuber, J.; Cummins, F. Learning to Forget: Continual Prediction with LSTM. In Proceedings of the 9th International Conference on Artificial Neural Networks, ICANN’99, Edinburgh, UK, 7–10 September 1999; pp. 850–855.

22. Ian, G.; Yoshua, B.; Aaron, C. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

23. Zhang, J.; Wei, Y.M.; Li, D.; Tan, Z.; Zhou, J. Short term electricity load forecasting using a hybrid model. *Energy* **2018**, *158*, 774–778. [CrossRef]

24. Xiong, T.; Bao, Y.; Hu, Z. Interval forecasting of electricity demand: A novel bivariate EMD-based support vector regression modeling framework. *Int. J. Electr. Power Energy Syst.* **2014**, *63*, 353–362. [CrossRef]

25. Lahouar, A.; Slama, J.B.H. Day-ahead load forecast using random forest and expert input selection. *Energy Convers. Manage.* **2015**, *103*, 1040–1051. [CrossRef]

26. Selakova, A.; Cvjetinovic, D.; Milovic, L.; Mellon, S.; Bekut, D. Hybrid PSO-SVM method for short-term load forecasting during periods with significant temperature variations in city of Burkhan. *Appl. Soft Comput.* **2014**, *16*, 80–88. [CrossRef]

27. Azim, H.; Davide, A.G.; Farshid, K.; Fabio, B.; Livio, D.S. Hybrid intelligent strategy for multifactor influenced electrical energy consumption forecasting. *Energy Sources Part B Econ. Plan. Policy* **2020**, *14*, 341–358.

28. Kavousi-fard, A.; Samet, H.; Marzbani, F. A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting. *Expert Syst.* **2014**, *41*, 6047–6056. [CrossRef]

29. Kim, T.; Cho, S. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* **2019**, *182*, 72–81. [CrossRef]

30. Zhang, J.; Tan, Z.; Wei, Y. An adaptive hybrid model for short term electricity price forecasting. *Appl. Energy* **2019**, *258*, 114087. [CrossRef]
31. Lago, J.; De Ridder, F.; De Schutter, B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Appl. Energy* **2018**, *221*, 386–405. [CrossRef]
32. Panapukidis, I.P.; Dagonas, A.S. Day-ahead electricity price forecasting via the application of artificial neural network based models. *Appl. Energy* **2016**, *172*, 132–151. [CrossRef]
33. Amjadiy, N.; Daraeepour, A.; Keynia, F. Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network. *IET Gener. Transm. Distrib.* **2010**, *4*, 432–444. [CrossRef]
34. Conejo, A.J.; Plazas, M.A.; Espinola, R.; Molina, A.B. Day-ahead electricity price forecasting using the wavelet transform and arima models. *IEEE Trans. Power Syst.* **2005**, *20*, 1035–1042. [CrossRef]
35. Amjadiy, N.; Keynia, F. Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. *IEEE Trans. Power Syst.* **2008**, *24*, 306–318. [CrossRef]
36. Luo, S.; Weng, Y. A two-stage supervised learning approach for electricity price forecasting by leveraging different data sources. *Appl. Energy* **2019**, *242*, 1497–1512. [CrossRef]
37. Yang, W.; Wang, J.; Niu, T.; Du, P. A hybrid forecasting system based on a dual decomposition strategy and multi-objective optimization for electricity price forecasting. *Appl. Energy* **2019**, *235*, 1205–1225. [CrossRef]
38. Amiri, M.; Davande, H.; Sadeghian, A.; Chartier, S. Feedback associative memory based on a new hybrid model of generalized regression and self-feedback neural networks. *Neural Netw.* **2010**, *23*, 892–904. [CrossRef]
39. Dabrowski, J.J.; Zhang, Y.; Rahman, A. ForecastNet: A Time-Variant Deep Feed-Forward Neural Network Architecture for ulti-step Ahead Time-Series Forecasting. In Proceedings of the International Conference on Neural Information Processing, Bangkok, Thailand, 18–22 November 2020; pp. 579–591.
40. Takeda, H.; Tamura, Y.; Sato, S. Using the ensemble Kalman filter for electricity load forecasting and analysis. *Energy* **2016**, *104*, 184–198. [CrossRef]
41. Chen, Y.; Zhang, S.; Zhang, W.; Peng, J.; Cai, Y. Multifactor spatio-temporal correlation model based on a combination of convolutional neural network and long short-term memory neural network for wind speed forecasting. *Energy Conver. Manage.* **2019**, *185*, 783–799. [CrossRef]
42. Ghedamsi, R.; Settou, N.; Gouareth, A.; Khamouli, A.; Saifi, N.; Reciou, B.; Dokkar, B. Modeling and forecasting energy consumption for residential buildings in Algeria using bottom-up approach. *Energy Build.* **2016**, *121*, 309–317. [CrossRef]
43. Dong, B.; Li, Z.; Rahman, S.M.; Vega, R. A hybrid model approach for forecasting future residential electricity consumption. *Energy Build.* **2016**, *117*, 341–351. [CrossRef]
44. Tascikaraoglu, A.; Sanandaji, B.M. Short-term residential electric load forecasting: A compressive spatial-temporal approach. *Energy Build.* **2016**, *111*, 380–392. [CrossRef]
45. Leung, M.C.; Norman, C.F.; Lai, L.L.; Chow, T.T. The use of occupancy space electrical power demand in building cooling load prediction. *Energy Build.* **2012**, *55*, 151–163. [CrossRef]
46. Yun, K.; Luck, R.; Mago, P.J.; Cho, H. Building hourly thermal load prediction using an indexed ARX model. *Energy Build* **2012**, *54*, 225–233. [CrossRef]
47. KEPCO’s iSMART. Available online: http://pcsc.kepco.co.kr/ (accessed on 13 March 2021).
48. YiRe, S.; Sanghoo, Y. Evaluation of weather information for electricity demand forecasting. *J. Korean Data Inf. Sci. Soc.* **2016**, *27*, 1601–1607.
49. Youjung, J.; Hyesook, H.; Baekjo, K.; Seongkyoun, K.; Giman, H.; Wookyun, L. The change of the average discomfort index from June to September during the past 10 years. *Clim. Chang. Res.* **2016**, *3*, 89–100.
50. Statistics KOREA Government Official Work Conference. Available online: https://www.index.go.kr/unify/idx-info.do?idCd=4291 (accessed on 4 August 2021).
51. Bengio, Y.L.; Courville, Y.; Vincent, P. Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [CrossRef]
52. Schmidhuber, J. Deep Learning in Neural Networks: An Overview. 2014. Available online: http://arxiv.org/abs/1404.7828 (accessed on 4 August 2021).
53. Szegedy, C.; Toshev, A.; Erhan, D. Deep neural networks for object detection. In Proceedings of the Advance Neural Information Processing Systems Conference, Lake Tahoe, NV, USA, 5–10 December 2013.
54. Dupond, S. A thorough review on the current advance of neural network structures. *Annu. Rev. Control* **2019**, *14*, 200–230.
55. Zell, A. Simulation Neuronaler Netze [Simulation of Neural Networks], 1st ed.; Addison-Wesley: Bonn, Germany, 1994. (In German)
56. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [CrossRef] [PubMed]
57. Cruse, H. *Neural Networks as Cybernetic Systems*, 2nd and revised ed.; Brains, Minds Media: Bielefeld, Germany, 2006.
58. Jaeger, H. Echo state network. *Scholarpedia* **2007**, *2*, 2330. [CrossRef]
59. Felix, A.G.; Schraudolph, N.N.; Jürgen, S. Learning Precise Timing with LSTM Recurrent Networks. *J. Mach. Learn. Res.* **2002**, *3*, 115–143.
60. Berglung, M.; Raiko, T.; Honkala, M.; Leo, K.; Vetek, A.; Karhunen, J. Bidirectional Recurrent Neural Networks as Generative Models. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 856–864.
61. Alex, S. Deriving the Recurrent Neural Network Definition and RNN Unrolling Using Signal Processing. In Proceedings of the Critiquing and Correcting Trends in Machine Learning Workshop at NeurIPS-2018, Montréal, QC, Canada, 3–8 December 2018.
62. Rainer, W.P.; Jun, T. How Hierarchical Control Self-organizes in Artificial Adaptive Systems. *Adapt. Behav.* 2005, 13, 211–225.
63. Giles, C.L.; Miller, C.B.; Chen, D.; Chen, H.H.; Sun, G.Z.; Lee, Y.C. Learning and Extracting Finite State Automata with Second-Order Recurrent Neural Networks. *Neural Comput.* 1992, 4, 393–405. [CrossRef]
64. Hochreiter, S. Untersuchungen zu dynamischen neuronalen Netzen. Diploma Thesis, Institut f. Informatik, Technische University Munich, Munich, Germany, 1991.
65. Hochreiter, S.; Bengio, Y.; Frasconi, P.; Schmidhuber, J. *Gradient Flow in Recurrent Nets: The Difficulty of Learning Long-Term Dependencies*; Wiley-IEEE Press: Hoboken, NJ, USA, 2001; pp. 237–243.
66. Valuva, M.V.; Nagornov, N.N.; Lyakhov, P.A.; Valuev, G.V.; Cheryvakov, N.I. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. In *Mathematics and Computers in Simulation*; Elsevier BV: Amsterdam, The Netherlands, 2020; Volume 177, pp. 232–243.
67. Mittelman, R. Time-series modeling with undecimated fully convolutional neural networks. *arXiv* 2015, arXiv:1508.00317.
68. Chen, Y.; Kang, Y.; Chen, Y.; Wang, Z. Probabilistic forecasting with temporal convolutional neural network. *Neurocomputing* 2009, 399, 491–501. [CrossRef]
69. Honglak, L.; Grosse, R.; Ranganath, R.; Ng, A.Y. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceedings of the 26th Annual International Conference on Machine Learning (ICML), Montreal, QC, Canada, 14–18 June 2009; pp. 609–616.
70. Hinton, G. Deep belief networks. *Scholarpedia* 2009, 4, 5947. [CrossRef]
71. Convolutional Deep Belief Networks on CIFAR-10. Available online: http://www.cs.toronto.edu/~kriz/conv-cifar10-aug2010.pdf (accessed on 5 August 2021).
72. Graves, A.; Fernández, S.; Gomez, F. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI), Orlando, FL, USA, 14–16 December 2006; pp. 369–376.
73. Wierstra, D.; Schmidhuber, J.; Gomez, F.J. Evolino: Hybrid Neuroevolution/Optimal linear search for sequence learning. In Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI), Edinburgh, Scotland, UK, 30 July–5 August 2005; pp. 853–858.
74. Afan, G.S.; Yaya, H.; Edi, A.; Wayan, S. Single layer & multilayer long short-term memory model with intermediate variables for weather forecasting. *Procedia Comput. Sci.* 2018, 135, 89–98.
75. Cho, K.; Van Merriënboer, B.; Bahdanau, D.; Bengio, Y. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. In Proceedings of the SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, Doha, Qatar, 25 October 2014; pp. 103–111.
76. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* 1997, 45, 2673–2681. [CrossRef]
77. Zhang, T.; Zhang, E.; Zhang, H.; Shen, F.; Guo, D.; Lin, K. An encoder-decoder neural network for indefinite length digit sequences in natural scene recognition. *J. Phys. Conf. Ser.* 2019, 1345, 022025. [CrossRef]
78. Seongchan, K.; Seungkyun, H.; Minsu, J.; Sakwang, S. Deep Rain: ConvLSTM network for precipitation prediction using multichannel radar data. *arXiv* 2017, arXiv:1711.02316.
79. Martin, L.; Zarzaledo, L.F.; Polo, J.; Navarro, A.; Marchante, R.; Cony, M. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. *Sol. Energy* 2010, 84, 1772–1781. [CrossRef]
80. Kleissel, J. *Solar Energy Forecasting and Resource Assessment*; Academic Press: Cambridge, MA, USA, 2013.
81. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *arXiv* 2015, arXiv:1412.6980.
82. Nair, V.; Hinton, G. Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010.
83. Agresti, A. *Categorical Data Analysis*; John Wiley and Sons: New York, NY, USA, 1990.
84. Willmott, C.; Matsuura, K. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in assessing average model performance. *Clim. Res.* 2005, 30, 79–82. [CrossRef]
85. Tensorflow.org. Deep Learning Library Developed by Google. Available online: https://www.tensorflow.org/ (accessed on 5 August 2021).
86. Keras.io. The Python Deep Learning Library. Available online: https://keras.io/ (accessed on 5 August 2021).