Article

Adaptive Optimized Pattern Extracting Algorithm for Forecasting Maximum Electrical Load Duration Using Random Sampling and Cumulative Slope Index

Jinseok Kim 1,2, Hyungseop Hong 1 and Ki-Il Kim 2,*

1 Department of KDN Electric power IT Research Institute, KEPCO KDN, Daejeon 34056, Korea; kim-668069@kdn.com (J.K.); ph.d.hong95@kdn.com (H.H.)
2 Department of Computer Science & Engineering, Chungnam National University, Daejeon 34134, Korea
* Correspondence: kikim@cnu.ac.kr; Tel.: +82-42-821-6856

Received: 23 May 2018; Accepted: 27 June 2018; Published: 1 July 2018

Abstract: Load forecasting techniques can be an essential method to save energy and shave peak loads in order to improve energy efficiency and maintain the stability of a power grid. To achieve this goal, machine learning-based approaches have been proposed recently. Before moving toward the long-term and ultimate solution such as machine learning, we propose a simple and efficient method to forecast electricity usage patterns and the duration of maximum electrical load using a small data set. The proposed algorithm can forecast maximum electrical load duration using random sampling and a cumulative slope index. To verify the algorithm, we utilized electricity data (from 2015.11 to 2016.12) obtained from a building with a constant lifestyle and electricity pattern. The performance of the algorithm was evaluated using electricity bills, the discharging condition of an energy storage system, and the cumulative slope index. It was found that the proposed algorithm could provide electricity cost savings of 0.62–2.28% compared with other, conventional electricity prediction techniques, such as the moving average method and exponential smoothing. In near future research, it is expected that this algorithm could be applied to electrical big data to handle real-time data processing.

Keywords: maximum electrical load duration; pattern forecasting; correlation analysis; random sampling; cumulative slope index; big data; real-time data processing

1. Introduction

There is a growing demand for electrical power as industry and economies develop that can sometimes result in insufficient power. To solve this problem, we are striving to expand the electricity supply facilities to enhance the power supply. However, despite our efforts, there is a limit to the extent to which supply facilities can be increased. Therefore, demand response (DR) management [1–4] of electrical power is a current trend and an alternative to focusing on expanding supply facilities. DR can provide a predictive method to control the power flow of an energy storage system (ESS), photovoltaic (PV) supply, buildings and grid [1] and based on a combination of a time-based and incentive-based DR program, a real-time incentive DR program can reduce the peak load [2]. DR programs using a game model [3,4] are effective in improving grid efficiency and reliability. The main policy of demand response is to improve energy efficiency and power load management. Regarding energy efficiency, forecasting electricity usage can be an essential method to save energy and shave peak loads. Research regarding load forecasting is currently in progress using statistical methods and time-series models.

Specifically, these load forecasting methods include regression analysis, exponential smoothing, and neural networks. Regression analysis [5–8] is a method used to predict the load and then utilize
the data to forecast load and independent variables such as temperature and GDP (Gross Domestic Product) [5]. Exponential smoothing models [9–12] offer an alternative method to forecast electricity usage using recently measured data and different types of weighting. This method can improve the error rate of predicted data because the recent data has a greater value in the weighting. On the other hand, older data has weight that decreases exponentially over time. Neural networks are one method to artificially model the human brain mathematically and such models can require more and different types of weights and variables. Neural networks can also be linked to non-linear models to distinguish between dependent variables and independent ones. The load forecasting method that combines soft computing and neural networks is the new approach called hybrid methodologies [13]. This model is different from traditional forecasting models such as statistical methods. Data mining framework can also be an intelligent analysis and prediction of energy consumption based on electricity data [14,15].

In this paper, we propose an algorithm for forecasting the duration of maximum electrical load using random sampling and a cumulative slope index. The reasons that we choose a proposed approach are because we do not need a lot of training data to forecast electricity usage patterns. Second, only electricity data (single variable) can be utilized to predict maximum electrical load durations. Finally, the proposed approach is a simple and efficient method to maintain prediction accuracy without a complicated model before moving toward a long-term and ultimate solution such as machine learning. What we propose to do here is to forecast maximum electrical load duration for a building with a constant lifestyle and electricity pattern on a weekly basis (except for holidays and weekends: Saturdays and Sundays). To verify the algorithm, we utilized the electricity data (2015.11–2016.12) obtained from the headquarters building of Korea Electric Power Knowledge Data & Network (KEPCO KDN) located in Naju City, South Korea. The reason that the proposed algorithm can forecast the future electrical load of buildings is because household or office buildings have relatively constant electrical usage patterns. Therefore, for these, we can forecast the pattern of electrical use and maximum electrical load duration using the algorithm. Normally, we might use conventional and sequential correlation analysis to forecast the pattern and load duration. However, the proposed algorithm is intended to forecast electrical use patterns and maximum electrical load duration using repetitive correlation analysis with data extracted from random sampling to analyze as many data as possible to model the pattern of electrical usage.

By applying the proposed approach, we offer the potential that the electrical use pattern can be obtained more quickly than by using sequential correlation analysis, in terms of a stochastic method. In the second section, we will see how conventional and sequential correlation analysis is used to extract an electrical use pattern that can represent monthly data. In the third section, we propose the algorithm, defining an alternative approach to forecast electrical use patterns that can represent monthly data, employing random sampling and repetitive correlation analysis. Finally, to evaluate and flesh out the performance of the proposed algorithm, we will forecast maximum electrical load durations with each of three models (random, moving average, and exponential smoothing model). Then, we will discuss the quantitative results of the proposed algorithm using random sampling.

2. Sequential Correlation Analysis for Extracting Representative Electrical Pattern

2.1. Sequential Correlation Analysis

Correlation involves the analysis of statistical and mathematical interrelations between components of large data sets. It can be utilized to analyze statistical interrelations between two probability variables in various fields of engineering. For example, we can study correlation of wind and hydro with electricity demand and prices [16] and develop a method to evaluate the characteristics
of team communications based on social network analysis [17]. To measure the correlation and similarity between two probability variables, covariance can be utilized as follows:

\[
\text{Cov}(X, Y) = \sigma_{XY} = E \{(X - \mu_X)(Y - \mu_Y)\} \\
= \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_X)(Y_i - \mu_Y) \\
= E(XY) - E(X)E(Y),
\]

(1)

The correlation coefficient can be defined as in the following equation:

\[
\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}
\]

(2)

The values of the correlation coefficient occur within the range \(-1\) to \(+1\) and when the correlation shows a linear relation, the correlation value gets closer to \(+1\). In contrast, when the correlation has a non-linear relationship, the value gets closer to \(-1\). Generally, if the correlation coefficient values are 0.7 or higher, they have a strong linear relationship. When the values are between 0.3 and 0.7, the linearity is moderate, and with values below 0.3, the correlation values have a weak linear relationship. Table 1 shows the statistical meanings of correlation coefficients.

| Correlation Value   | Meaning                        |
|---------------------|--------------------------------|
| \(-1\) to \(-0.7\) | Strong negative linear relationship |
| \(-0.7\) to \(-0.3\) | Moderate negative linear relationship |
| \(-0.1\) to \(+0.1\) | Negligible linear relationship |
| \(+0.1\) to \(+0.3\) | Weak positive linear relationship |
| \(+0.3\) to \(+0.7\) | Moderate positive linear relationship |
| \(+0.7\) to \(+1.0\) | Strong positive linear relationship |

In this paper, we apply a correlation coefficient threshold of 0.8 (or 0.7) to decide the representative electrical pattern. This means that the electrical use pattern is defined as the pattern with correlation coefficient \(\geq 0.7\), which means a strong positive association between patterns.

2.2. Method for Extracting Representative Electrical Patterns for Each Month from the Original Data

In this section, we address the method used to extract the representative electrical pattern for each month from original data for the KEPCO KDN headquarters. The representative electrical pattern means the pattern of electrical usage of a specific date each month that could represent usage during the month. The main reason to extract the representative electrical pattern is to examine the performance of forecasting maximum electrical load duration for each model (random, moving average, and exponential smoothing model) compared to a representative electrical pattern. If the electrical pattern of a specific date within each month has probabilistically obvious affinities with the patterns of other dates (correlation coefficient \(\geq\) correlation coefficient threshold) and also is in the majority within each month, the pattern can be a representative electrical pattern of the month. In this paper, a representative electrical pattern is defined as the pattern with correlation coefficient threshold of 0.8 that accounts for \(\geq 65\%\) (ratio threshold) within the data for a month. However, if any pattern within a month does not satisfy that condition, we apply the second definition of representative electrical pattern with correlation coefficient threshold of 0.7 that accounts for \(\geq 65\%\) within the data for a month. To achieve a representative electrical pattern, we perform correlation analysis sequentially using pre-processed data extracted from the original electricity data.

The Algorithm 1 explains how to extract the representative electrical pattern using sequential correlation analysis. As shown in Algorithm 1, to achieve a representative electrical pattern, we perform
correlation analysis sequentially using pre-processed data extracted from the original electricity data. At first the correlation coefficient threshold was set at the correlation coefficient of 0.8, and then the ratio threshold was set at 65%. Second, we performed correlation analysis between components of two data sets sequentially. Finally, the pattern with the highest ratio among electrical patterns with a correlation coefficient $\geq 0.8$ that accounts for $\geq 65\%$ was extracted. As already mentioned, if any pattern within a month does not satisfy that condition, we apply the second definition of representative electrical pattern with correlation coefficient $\geq 0.7$ that accounts for $\geq 65\%$ within the data for a month, and then execute algorithm 1 once again.

**Algorithm 1: The algorithm to extract the representative electrical pattern**

1: Cut_correl $\leftarrow$ Correlation coefficient threshold  
2: Cut_ratio $\leftarrow$ Ratio threshold  
3: Correl($i,j$) $\leftarrow$ Statistical correlation function between $i$–th dataset and $j$–th dataset  
4: $n$ $\leftarrow$ The last numberfield among dataset $\{\text{Date}_{1..n}, \text{Electrical Load}_{1..n}\}$  
5: Record$_i$ = $\{T_i, R_i, \text{Date}_i, \text{Electrical Load}_i\}$  
6: $B$ $\leftarrow$ Memory structure to store Record$_i$  
7: while $i \leq n$ do  
8: while $j \leq n$ do  
9: $C_{i,j} \leftarrow \text{Correl}(i,j)$  
10: if ($C_{i,j} \geq \text{Cut\_correl}$) then  
11: $P_{i,j} \leftarrow 1$  
12: else  
13: $P_{i,j} \leftarrow 0$  
14: end if  
15: end while  
16: $\text{Ratio}_i = \frac{\sum_{j=1}^{n} P_{i,j}}{n}$  
17: if ($\text{Ratio}_i \geq \text{Cut\_ratio}$) then  
18: $B \leftarrow B \cup \text{Record}_i$  
19: end if  
20: end while  
21: $k \leftarrow \text{argmax}_{p \in B} (\text{Record}_p \cdot T_p)$  
22: Extracted_Pattern $\leftarrow \text{Record}_k \cdot \text{Electrical Load}_k$

The concept of extracting representative electrical patterns and maximum electrical load duration using sequential correlation analysis is presented in Figure 1.

To forecast the duration of the maximum electrical load, we calculated cumulative slope index values for a representative electrical pattern and then, using the index, we compared it with each model for quantitative evaluation. The ratio of the representative electrical pattern with correlation coefficient $\geq 0.8$ (or $\geq 0.7$) is listed in Table 2. All these values were higher than 65% in case of correlation coefficient $\geq 0.8$ (or $\geq 0.7$) and represented the majority within each month. From these results, the use pattern of each month can be a representative of a statistical model of each month’s data in terms of stochastic method. As shown in Figures 2–7, the representative electrical patterns of a specific date within each month are different from each other. From these results, we could roughly know how the maximum electrical load durations of representative electrical patterns that we want to forecast are also different from each other.
Figure 1. Process for extracting representative electrical patterns and maximum electrical load duration using sequential correlation analysis.

Table 2. Ratios of representative electrical patterns with correlation coefficient ≥0.8 (or ≥0.7).

| Representative Electrical Pattern | Ratio (%) (Correlation Coefficient ≥ 0.8) | Ratio (%) (Correlation Coefficient ≥ 0.7) |
|----------------------------------|------------------------------------------|------------------------------------------|
| 2016.01                          | 50                                       | 70                                       |
| 2016.02                          | 48                                       | 67                                       |
| 2016.03                          | 70                                       | 83                                       |
| 2016.04                          | 86                                       | 90                                       |
| 2016.05                          | 95                                       | 100                                      |
| 2016.06                          | 91                                       | 95                                       |
| 2016.07                          | 100                                      | 100                                      |
| 2016.08                          | 96                                       | 96                                       |
| 2016.09                          | 86                                       | 86                                       |
| 2016.10                          | 76                                       | 86                                       |
| 2016.11                          | 68                                       | 77                                       |
| 2016.12                          | 73                                       | 82                                       |
Figure 2. Representative pattern of electrical usage. (a) (2016.01); (b) (2016.02).

Figure 3. Representative pattern of electrical usage. (a) (2016.03); (b) (2016.04).

Figure 4. Representative pattern of electrical usage. (a) (2016.05); (b) (2016.06).
Figure 5. Representative pattern of electrical usage. (a) (2016.07); (b) (2016.08).

Figure 6. Representative pattern of electrical usage. (a) (2016.09); (b) (2016.10).

Figure 7. Representative pattern of electrical usage. (a) (2016.11); (b) (2016.12).
3. Proposed Algorithm Using Random Sampling

In this section, we address the proposed algorithm using random sampling and repeated correlation analysis. Using the random sampling technique and repeated correlation analysis with previous data (such as one month ago or two months ago) within the period 2015.11–2016.12, we were able to find electricity usage patterns with strong correlation value in the electric data set. Actually, random sampling or Monte Carlo techniques [18–21] are widely used to solve various kinds of problems through the computation of random numbers. This research usually deals with uncertainty and the sensitivity of the energy model. These are also used to provide statistical estimates in various fields of engineering. These techniques can be applied to evaluate power systems’ composite reliability [18,19] and used for sensitivity analysis of scenario ranking for energy generation [20]. With the growing use of renewable energy resources, distributed generation (DG) systems are spreading and Monte Carlo techniques can address the probabilistic assessment of DG penetration to the distribution network [21]. Compared with sequential correlation analysis, the advantage of a correlation using random sampling is that it offers the potential to obtain electrical use patterns more quickly than with sequential correlation analysis, in terms of stochastic method. Using the proposed algorithm and random sampling, we might do less repetitive correlation analysis by using only the data extracted by random sampling. Algorithm 2 explains how to extract the pattern using random sampling. We perform correlation analysis between components of two data sets using random sampling to achieve a predicted pattern. At first, the correlation coefficient threshold was set at the correlation coefficient of 0.8, and then the ratio threshold was set at the specific value we want to achieve. The maximum repeated count was set at M values. Second, we performed correlation analysis between components of two data sets randomly and repeatedly. If the electrical pattern of a specific date was above the ratio threshold, the electrical use pattern for that sample date was recorded. Finally, when we encountered the maximum repeated count, the pattern among the sample dates with the highest ratio was extracted.

Algorithm 2: Proposed algorithm using random sampling

1: Cut_correl ← Correlation coefficient threshold
2: Cut_ratio ← Ratio threshold
3: Correl(i, R) ← Statistical correlation function between i-th dataset and R-th dataset
4: M ← Maximum repeated count
5: n ← The last numberfield among datasets \{Date_{(1..n)}, Electrical_Load_{(1..n)}\}
6: Record_i = \{T_i, R_i, Date_i, Electrical_Load_i\}
7: B ← Memory structure to store Record_i
8: while count ≤ M do
9: R ← Random Variable(1..n)
10: while i ≤ n do
11: C_{i,R} ← Correl(i, R)
12: if (C_{i,R} ≥ Cut_correl) then
13: P_{i,R} ← 1
14: else
15: P_{i,R} ← 0
16: end if
17: end while
18: Ratio_count = \sum_{i=1}^{n} \frac{P_{i,R}}{R}
19: if (Ratio_count ≥ Cut_ratio) then
20: B ← B \cup Record_count
21: end if
The concept of extracting the predicted pattern using random sampling is depicted in Figure 8.

To verify the appropriateness of the algorithm, we compared each representative electrical pattern with a predicted pattern extracted by the proposed algorithm. Comparisons of representative and predicted patterns are shown in Figures 9–14. As shown in those figures, most of the predicted patterns...
bear a resemblance to the representative electrical pattern. From these results, we could roughly see the proposed algorithm can predict the maximum electrical load duration of the representative electrical pattern. However, the predicted patterns in Figure 14 cannot be seen as analogous to the representative electrical patterns. The cause of the problem might be that the link between the constant lifestyle and electrical load pattern are not the same as for the other patterns. In the next section, to evaluate the quantitative performance of the proposed algorithm using random sampling, we discuss the prediction results of maximum electrical load duration.

Figure 9. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.01) and the predicted pattern from 2015.11–2015.12; (b) REC (2016.02) and the predicted pattern from 2015.12–2016.01.

Figure 10. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.03) and the predicted pattern from 2016.01–2016.02; (b) REC (2016.04) and the predicted pattern from 2016.02–2016.03.
Figure 9. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.01) and the predicted pattern from 2015.11–2015.12; (b) REC (2016.02) and the predicted pattern from 2015.12–2016.01.

Figure 10. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.03) and the predicted pattern from 2016.01–2016.02; (b) REC (2016.04) and the predicted pattern from 2016.02–2016.03.

Figure 11. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.05) and the predicted pattern from 2016.03–2016.04; (b) REC (2016.06) and the predicted pattern from 2016.04–2016.05.

Figure 12. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.07) and the predicted pattern from 2016.05–2016.06; (b) REC (2016.08) and the predicted pattern from 2016.06–2016.07.

Figure 13. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.09) and the predicted pattern from 2016.07–2016.08; (b) REC (2016.10) and the predicted pattern from 2016.08–2016.09.

Figure 14. Comparison between a representative electrical pattern (REC) and the pattern predicted from randomly sampled data. (a) REC (2016.11) and the predicted pattern from 2016.09–2016.10; (b) REC (2016.12) and the predicted pattern from 2016.10–2016.11.
4. Performance Results of the Proposed Algorithm

In this section, we examine the performance of the proposed algorithm. For evaluation of its performance, we discuss the results of maximum electrical load duration. The reason that we attempted to find the maximum electrical load duration was to examine the electricity cost saving effects from use of the proposed algorithm. Maximized electricity cost saving effects occur when ESS are discharging electricity during the period of maximum electrical load. To confirm the duration of maximum electrical load, we propose using the concept of cumulative slope index, as well as the ESS discharging condition and KEPCO electricity bills, to evaluate a quantitative analysis. The flowchart for evaluating the performance is shown in Figure 15. As shown in Figure 15, we extracted pre-processed data from the original data and performed correlation analysis and then calculated the cumulative slope index values for each model. Finally, we compared them with the cumulative slope index value of representative electrical loads for quantitative evaluation.
4. Performance Results of the Proposed Algorithm

In this section, we examine the performance of the proposed algorithm. For evaluation of its performance, we discuss the results of maximum electrical load duration. The reason that we attempted to find the maximum electrical load duration was to examine the electricity cost saving effects from use of the proposed algorithm. Maximized electricity cost saving effects occur when ESS are discharging electricity during the period of maximum electrical load. To confirm the duration of maximum electrical load, we propose using the concept of cumulative slope index, as well as the ESS discharging condition and KEPCO electricity bills, to evaluate a quantitative analysis. The flowchart for evaluating the performance is shown in Figure 15. As shown in Figure 15, we extracted pre-processed data from the original data and performed correlation analysis and then calculated the cumulative slope index values for each model. Finally, we compared them with the cumulative slope index value of representative electrical loads for quantitative evaluation.

4.1. The Concept of Cumulative Slope Index, Energy Storage System (ESS) Discharging Condition, and Korea Electric Power (KEPCO) Electricity Bill

4.1.1. The Concept of Cumulative Slope Index

To evaluate the algorithm quantitatively, the concept of the cumulative slope index can be used. The definition of the cumulative slope index, including other parameters, is shown in Table 3.
Table 3. Parameter list of time, power, slope, cumulative slope and cumulative slope index.

| Parameter                                             | Parameter | Parameter | Parameter | Parameter |
|-------------------------------------------------------|-----------|-----------|-----------|-----------|
| Time [h]                                               | $T_0$     | $T_1$     | $T_2$     | $T_n$     |
| Power of Predicted Pattern [kWh]                      | $P_0$     | $P_1$     | $P_2$     | $P_n$     |
| Power of Representative Electrical Pattern [kWh]      | $N_0$     | $N_1$     | $N_2$     | $N_n$     |
| Slope of Predicted Pattern                            | $X_0$     | $X_1$     | $X_2$     | $X_{n-1}$ |
| Slope of Representative Electrical Pattern             | $S_0$     | $S_1$     | $S_2$     | $S_{n-1}$ |
| Cumulative Slope of Predicted Pattern $[\%]$          | $C_{X0}$  | $C_{X1}$  | $C_{X2}$  | $C_{X_{n-1}}$ |
| Cumulative Slope of Representative Electrical Pattern $[\%]$ | $C_{S0}$  | $C_{S1}$  | $C_{S2}$  | $C_{S_{n-1}}$ |
| Cumulative Slope Index of Predicted Pattern $[\%]$    | $C_{P0}$  | $C_{P1}$  | $C_{P2}$  | $C_{P_{n-1}}$ |
| Cumulative Slope Index of Representative Electrical Pattern $[\%]$ | $C_{R0}$  | $C_{R1}$  | $C_{R2}$  | $C_{R_{n-1}}$ |

The slope of the predicted pattern is determined as follows:

$$X_{n-1} = \frac{P_n - P_{n-1}}{T_n - T_{n-1}}$$  \hspace{1cm} (3)

The slope of the representative electrical pattern is calculated as follows:

$$S_{n-1} = \frac{N_n - N_{n-1}}{T_n - T_{n-1}}$$  \hspace{1cm} (4)

The cumulative slope index of the predicted pattern is calculated as follows:

$$T_0 : X_0 = C_{X0} \hspace{2cm} C_{P0} = \frac{C_{X0}}{C_{X_{\text{max}}}} \times 100[\%]$$
$$T_1 : X_0 + X_1 = C_{X1} \hspace{2cm} C_{P1} = \frac{C_{X1}}{C_{X_{\text{max}}}} \times 100[\%]$$
\vdots
\begin{equation}
T_{n-1} : X_{n-2} + X_{n-1} = C_{X_{n-1}} \hspace{2cm} C_{P_{n-1}} = \frac{C_{X_{n-1}}}{C_{X_{\text{max}}}} \times 100[\%],
\end{equation}

The cumulative slope index of the representative electrical pattern is calculated as follows:

$$T_0 : S_0 = C_{S0} \hspace{2cm} C_{R0} = \frac{C_{S0}}{C_{S_{\text{max}}}} \times 100[\%]$$
$$T_1 : S_0 + S_1 = C_{S1} \hspace{2cm} C_{R1} = \frac{C_{S1}}{C_{S_{\text{max}}}} \times 100[\%]$$
\vdots
\begin{equation}
T_{n-1} : S_{n-2} + S_{n-1} = C_{S_{n-1}} \hspace{2cm} C_{R_{n-1}} = \frac{C_{S_{n-1}}}{C_{S_{\text{max}}}} \times 100[\%],
\end{equation}

The slope ($X$, $S$ parameter) is defined as the electricity power variation divided by the time change. The cumulative slope of predicted pattern ($C_X$ parameter) and the cumulative slope of representative electrical pattern ($C_S$ parameter) are defined as the sum of the next slope value and the cumulative slope value. The cumulative slope index ($C_P$, $C_R$ parameter) is defined as the ratio of the cumulative slope divided by the maximum cumulative slope ($C_{X_{\text{max}}}$, $C_{S_{\text{max}}}$ parameter). For example, if the cumulative slope index is $<80\%$, it means that the load is not within the maximum electrical load duration. In contrast, when the cumulative slope index is $\geq 80\%$, it indicates times within the maximum electrical load duration.

4.1.2. ESS Discharging Condition

Normally, an ESS cannot discharge continuously because of its limited capacity and lifecycle. Considering these characteristics, it is efficient to discharging an ESS only during the period of maximum electrical load. Therefore, the ESS is in discharging condition only while the cumulative
slope index is over 80% (i.e., which means the electrical usage is 80% of maximum load). In this paper, the maximum electrical load duration is defined as being when the cumulative index is over 80%. When the ESS is discharging during the maximum electrical load, utility electricity is not used. At that time, we can provide electricity cost saving effects compared to using utility energy all the time. We evaluated the performance of the proposed algorithm under the ESS discharging condition that the cumulative slope index was over 80%.

4.1.3. KEPCO Electricity Bills

The KEPCO electricity bills were used to evaluate a quantitative analysis. The electricity costs from the bills are shown in Tables 4 and 5. We were able to perform quantitative evaluation by comparing the energy use predicted using the proposed algorithm against the predictions made with other techniques, using the actual costs from the bills.

| Table 4. Korea Electric Power (KEPCO) electricity bill of high voltage (A) type. |
|----------------------------------|
| **High Voltage (A)** | **Basic Charge (KRW/kW)** | **Time Zone** | **Electricity Usage Fare (KRW/kWh)** |
| | | | **Summer** | **Spring, Autumn** | **Winter** |
| | | | **June–August** | **March–May, Sep.–Oct.** | **Nov.–Feb.** |
| 7220 | Low load | 61.6 | 61.6 | 68.6 |
| | Medium load | 114.5 | 84.1 | 114.7 |
| | Maximum load | 196.6 | 114.8 | 172.2 |

| Table 5. KEPCO electricity bill for time zone and season. |
|----------------------------------|
| **Time Zone** | **Season** |
| | **Summer** | **Spring, Autumn** | **Winter** |
| | **June–August** | **March–May, Sep.–Oct.** | **Nov.–Feb.** |
| Low load | 23:00–09:00 | 23:00–09:00 | 23:00–09:00 |
| Medium load | 09:00–10:00 | 09:00–10:00 | 09:00–10:00 |
| | 12:00–13:00 | 12:00–13:00 | 12:00–17:00 |
| | 17:00–23:00 | 17:00–23:00 | 20:00–22:00 |
| Maximum load | 10:00–12:00 | 10:00–12:00 | 10:00–12:00 |
| | 13:00–17:00 | 13:00–17:00 | 17:00–20:00 |
| | | | 22:00–23:00 |

4.2. Other Conventional Electricity Prediction Models (Moving Average Method and Exponential Smoothing)

Two conventional electricity prediction techniques are introduced here for comparison. One is the moving average method and the other is the exponential smoothing model.

The moving average is defined as follows:

\[ F_t = \frac{1}{n} (D_{t-1} + D_{t-2} + \cdots + D_{t-n}) \quad (7) \]

As a sort of smoothing time series, this method is widely used to forecast the electricity usage because the model is simple and practical. By applying the model, we can extract the predicted pattern using the moving average method from average previous data (one and two months ago). The averaged previous data and the predicted pattern extracted from the averaged previous data using the moving average method are summarized in Table 6.
Table 6. Average previous data and the predicted pattern extracted from averaged previous data using the moving average method.

| Averaged Data for | Predicted Pattern for |
|------------------|-----------------------|
| 2015.11–12       | 2016.01               |
| 2015.12–2016.01  | 2016.02               |
| 2016.01–02       | 2016.03               |
| 2016.02–03       | 2016.04               |
| 2016.03–04       | 2016.05               |
| 2016.04–05       | 2016.06               |
| 2016.05–06       | 2016.07               |
| 2016.06–07       | 2016.08               |
| 2016.07–08       | 2016.09               |
| 2016.08–09       | 2016.10               |
| 2016.09–10       | 2016.11               |
| 2016.10–11       | 2016.12               |

The exponential smoothing model is another sort of smoothing time series. However, there are some differences in the data size used, for forecasting using the moving average method and the exponential smoothing model. While the moving average method needs all the previous data (one and two months ago), the exponential smoothing model needs only a small amount of recent data.

In this paper, we applied data from the previous week (5 days) to forecast the monthly pattern. For example, to forecast the pattern of the month 2016.03, we applied the data from five days (02/23, 02/24, 02/25, 02/26, and 02/29) to forecast the pattern. The model is defined as follows:

$$F_t = \alpha X_{t-1} + \alpha (1 - \alpha) X_{t-2} + \alpha (1 - \alpha)^2 X_{t-3} + \alpha (1 - \alpha)^3 X_{t-4} + \cdots + \alpha (1 - \alpha)^n X_{t-n}$$  \hspace{1cm} (8)

As you can see from the equation, the most recent data gets the largest weighting and the old data has weights that decrease exponentially over time. Each date having a different weight for forecasting a monthly pattern is listed in Table 7.

Table 7. The predicted pattern list extracted from the exponential smoothing model using different weights of each date.

| Previous Data | Predicted Pattern |
|---------------|-------------------|
| Weight (\(\alpha\)) | - |
| 0.5 | 0.25 | 0.125 | 0.0625 | 0.03125 | - |
| 12/31 | 12/30 | 12/29 | 12/28 | 12/25 | 2016.01 |
| 1/29 | 1/28 | 1/27 | 1/26 | 1/25 | 2016.02 |
| 2/29 | 2/26 | 2/25 | 2/24 | 2/23 | 2016.03 |
| 3/31 | 3/30 | 3/29 | 3/28 | 3/25 | 2016.04 |
| 4/29 | 4/28 | 4/27 | 4/26 | 4/25 | 2016.05 |
| 5/31 | 5/30 | 5/27 | 5/26 | 5/25 | 2016.06 |
| 6/30 | 6/29 | 6/28 | 6/27 | 6/24 | 2016.07 |
| 7/29 | 7/28 | 7/27 | 7/26 | 7/25 | 2016.08 |
| 8/31 | 8/30 | 8/29 | 8/26 | 8/25 | 2016.09 |
| 9/30 | 9/29 | 9/28 | 9/27 | 9/26 | 2016.10 |
| 10/31 | 10/28 | 10/27 | 10/26 | 10/25 | 2016.11 |
| 11/30 | 11/29 | 11/28 | 11/25 | 11/24 | 2016.12 |

4.3. The Performance Results of the Proposed Algorithm

In this section, we examine the performance results of the proposed algorithm. To evaluate the performance, we address the ability to forecast the maximum electrical duration using the KEPCO electricity bills, ESS discharging condition, and cumulative slope index.
As already mentioned, the proposed algorithm performs correlation analysis between components of two data sets using random sampling to achieve a predicted pattern. Using the proposed algorithm, we might do less repetitive correlation analysis by using only the data extracted by random sampling. Then, we compare the performance results of the proposed algorithm with other, more conventional electricity prediction techniques such as moving average and exponential smoothing. To evaluate the performance of three models (random sampling, moving average, and exponential smoothing), the evaluation process was divided into two parts: forecasting accuracy and cost savings.

According to the predicted accuracy for quantitative evaluation, we could use cumulative slope index values. If the cumulative slope index of the representative electrical load was <80%, the load was not at maximum. In contrast, when the cumulative slope index of the representative electrical load was ≥80%, it means that conditions were within the maximum electrical load duration. The results of the cumulative slope index values for each model and representative electrical load are shown in Tables 8–19.

Table 8. Cumulative slope index values for three models (2015.11–12) and base model (2016.01).

| Model Group   | Cumulative Slope Index | Maximum Load Time Zone | CX_{max} |
|---------------|------------------------|------------------------|----------|
|               | 10–11                  | 11–12                  | 17–18    | 18–19 | 19–20 | 22–23 |
| Random        | CX                     | 148.376                | 115.234  | 112.366 | 79.101 | −7.812 | −178.22 | 159.973 |
|               | CP                     | 93%                    | 72%      | 70%   | 49%    | −5%    | −111%   | -     |
| Moving Average| CX                     | 240.631                | 219.31   | 209.799 | 128.534 | 52.983 | −78.845 | 250.356 |
|               | CP                     | 96%                    | 88%      | 84%   | 51%    | 21%    | −30%    | -     |
| Exponential Smoothing | CX | 354.525          | 337.624  | 227.709 | 101.349 | 44.494 | −55.237 | 354.525 |
|               | CP                     | 100%                   | 95%      | 64%   | 29%    | 13%    | −16%    | -     |
| Base Model    | Cumulative Slope Index | Maximum Load Time Zone | CS_{max} |
|               | 10–11                  | 11–12                  | 17–18    | 18–19 | 19–20 | 22–23 |
| Representative| CS                     | 226.318                | 166.26   | 85.449 | −24.902 | −138.67 | −255.37 | 288.696 |
| Electrical Load (2016.01) | CR | 78%                    | 58%      | 30%   | −9%    | −48%   | −88%    | -     |

Table 9. Cumulative slope index values for three models (2015.12–2016.01) and base model (2016.02).

| Model Group   | Cumulative Slope Index | Maximum Load Time Zone | CX_{max} |
|---------------|------------------------|------------------------|----------|
|               | 10–11                  | 11–12                  | 17–18    | 18–19 | 19–20 | 22–23 |
| Random        | CX                     | 382.568                | 385.923  | 217.346 | 152.832 | 121.582 | −108.89 | 385.925 |
|               | CP                     | 99%                    | 100%     | 50%   | 40%    | 32%    | −28%    | -     |
| Moving Average| CX                     | 256.708                | 231.616  | 211.072 | 113.914 | 22.538  | −129.79 | 256.708 |
|               | CP                     | 100%                   | 90%      | 82%   | 44%    | 9%     | −51%    | -     |
| Exponential Smoothing | CX | 125.557          | 95.304   | 52.869  | −32.944 | −127.27 | −255.57 | 160.351 |
|               | CP                     | 78%                    | 59%      | 33%   | −21%   | −79%   | −159%   | -     |
| Base Model    | Cumulative Slope Index | Maximum Load Time Zone | CS_{max} |
|               | 10–11                  | 11–12                  | 17–18    | 18–19 | 19–20 | 22–23 |
| Representative| CS                     | 407.166                | 365.845  | 296.265 | 262.695 | 124.512 | 59.082  | 481.018 |
| Electrical Load (2016.02) | CR | 85%                    | 76%      | 62%   | 55%    | 26%    | 12%     | -     |
Table 10. Cumulative slope index values for three models (2016.01–02) and base model (2016.03).

| Model Group            | Cumulative Slope Index | Maximum Load Time Zone | CX_{max} |
|------------------------|------------------------|------------------------|----------|
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Random                 | CX                     | 347.717                | 353.638  | 346.558  | 330.566  | 320.007  | 316.833  | 578.064  |
|                        | CP                     | 60%                    | 61%      | 60%      | 57%      | 55%      | 55%      | -        |
| Moving Average         | CX                     | 236.716                | 210.827  | 200.841  | 174.209  | 175.091  | 170.832  | 252.427  |
|                        | CP                     | 94%                    | 84%      | 80%      | 69%      | 69%      | 68%      | -        |
| Exponential Smoothing  | CX                     | 297.642                | 278.204  | 274.313  | 254.744  | 222.309  | 246.555  | 299.379  |
|                        | CP                     | 99%                    | 93%      | 92%      | 85%      | 74%      | 82%      | -        |
| Base Model             | Cumulative Slope Index | Maximum Load Time Zone | CS_{max} |
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Representative Load    | CS                     | 167.297                | 154.785  | 86.853   | 49.6826  | 64.5752  | 106.873  | 219.482  |
|                        | CR                     | 76%                    | 71%      | 40%      | 23%      | 29%      | 49%      | -        |

Table 11. Cumulative slope index values for three models (2016.02–03) and base model (2016.04).

| Model Group            | Cumulative Slope Index | Maximum Load Time Zone | CX_{max} |
|------------------------|------------------------|------------------------|----------|
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Random                 | CX                     | 257.629                | 237.915  | 250.854  | 232.056  | 195.557  | 177.185  | 279.602  |
|                        | CP                     | 92%                    | 85%      | 90%      | 83%      | 70%      | 63%      | -        |
| Moving Average         | CX                     | 245.565                | 219.137  | 214.525  | 183.964  | 190.482  | 178.725  | 284.44   |
|                        | CP                     | 86%                    | 77%      | 75%      | 65%      | 67%      | 63%      | -        |
| Exponential Smoothing  | CX                     | 265.12                 | 262.997  | 255.043  | 249.754  | 249.271  | 219.193  | 332.859  |
|                        | CP                     | 80%                    | 79%      | 77%      | 75%      | 75%      | 66%      | -        |
| Base Model             | Cumulative Slope Index | Maximum Load Time Zone | CS_{max} |
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Representative Load    | CS                     | 237.427                | 217.224  | 226.074  | 219.421  | 250.977  | 216.187  | 250.97   |
|                        | CR                     | 95%                    | 87%      | 90%      | 87%      | 100%     | 86%      | -        |

Table 12. Cumulative slope index values for three models (2016.03–04) and base model (2016.05).

| Model Group            | Cumulative Slope Index | Maximum Load Time Zone | CX_{max} |
|------------------------|------------------------|------------------------|----------|
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Random                 | CX                     | 281.128                | 272.217  | 316.04   | 300.476  | 292.847  | 252.93   | 316.04   |
|                        | CP                     | 89%                    | 86%      | 100%     | 95%      | 95%      | 80%      | -        |
| Moving Average         | CX                     | 258.72                 | 237.866  | 247.406  | 258.707  | 222.408  | 217.676  | 283.673  |
|                        | CP                     | 91%                    | 84%      | 87%      | 91%      | 78%      | 77%      | -        |
| Exponential Smoothing  | CX                     | 255.423                | 229.015  | 226.505  | 299.784  | 152.504  | 188.19   | 299.784  |
|                        | CP                     | 85%                    | 76%      | 76%      | 100%     | 51%      | 63%      | -        |
| Base Model             | Cumulative Slope Index | Maximum Load Time Zone | CS_{max} |
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Representative Load    | CS                     | 323.486                | 338.135  | 340.698  | 348.389  | 350.586  | 332.397  | 395.386  |
|                        | CR                     | 82%                    | 86%      | 86%      | 88%      | 89%      | 84%      | -        |

Table 13. Cumulative slope index values for three models (2016.04–05) and base model (2016.06).

| Model Group            | Cumulative Slope Index | Maximum Load Time Zone | CX_{max} |
|------------------------|------------------------|------------------------|----------|
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Random                 | CX                     | 256.287                | 253.845  | 249.451  | 221.497  | 223.511  | 204.285  | 256.287  |
|                        | CP                     | 100%                   | 99%      | 97%      | 86%      | 87%      | 80%      | -        |
| Moving Average         | CX                     | 262.594                | 295.04   | 297.768  | 311.254  | 280.826  | 291.83   | 311.254  |
|                        | CP                     | 91%                    | 94%      | 96%      | 100%     | 90%      | 94%      | -        |
| Exponential Smoothing  | CX                     | 380.074                | 404.259  | 459.457  | 424.423  | 424.335  | 431.255  | 459.457  |
|                        | CP                     | 83%                    | 88%      | 100%     | 92%      | 94%      | 94%      | -        |
| Base Model             | Cumulative Slope Index | Maximum Load Time Zone | CS_{max} |
|                        | 10–11                  | 11–12                  | 13–14    | 14–15    | 15–16    | 16–17    |          |
| Representative Load    | CS                     | 561.035                | 545.776  | 576.05   | 597.656  | 595.384  | 605.713  | 605.713  |
|                        | CR                     | 93%                    | 90%      | 95%      | 99%      | 98%      | 100%     | -        |
### Table 14. Cumulative slope index values for three models (2016.05–06) and base model (2016.07).

| Model Group      | Cumulative Slope Index | Maximum Load Time Zone | CX<sub>max</sub> |
|------------------|------------------------|------------------------|------------------|
|                  | 10–11  | 11–12  | 13–14 | 14–15 | 15–16 | 16–17 |
| Random           | CX 515.747 | 412.598 | 471.558 | 463.745 | 365.845 | 515.747 |
|                  | CP  | 89% | 100% | 80% | 91% | 90% | 71% |
| Moving Average   | CX 378.898 | 382.369 | 345.584 | 347.651 | 338.853 | 382.369 |
|                  | CP  | 95% | 99% | 100% | 90% | 91% | 89% |
| Exponential Smoothing | CX 471.558 | 497.798 | 256.737 | 237.038 | 208.317 | 452.591 |
|                  | CP  | 100% | 93% | 97% | 57% | 52% | 46% |

### Table 15. Cumulative slope index values for three models (2016.06–07) and base model (2016.08).

| Model Group      | Cumulative Slope Index | Maximum Load Time Zone | CX<sub>max</sub> |
|------------------|------------------------|------------------------|------------------|
|                  | 10–11  | 11–12  | 13–14 | 14–15 | 15–16 | 16–17 |
| Random           | CX 515.747 | 412.598 | 471.558 | 463.745 | 365.845 | 515.747 |
|                  | CP  | 89% | 100% | 80% | 91% | 90% | 71% |
| Moving Average   | CX 345.584 | 382.369 | 345.584 | 347.651 | 338.853 | 382.369 |
|                  | CP  | 95% | 99% | 100% | 90% | 91% | 89% |
| Exponential Smoothing | CX 471.558 | 497.798 | 256.737 | 237.038 | 208.317 | 452.591 |
|                  | CP  | 100% | 93% | 97% | 57% | 52% | 46% |

### Table 16. Cumulative slope index values for three models (2016.07–08) and base model (2016.09).

| Model Group      | Cumulative Slope Index | Maximum Load Time Zone | CX<sub>max</sub> |
|------------------|------------------------|------------------------|------------------|
|                  | 10–11  | 11–12  | 13–14 | 14–15 | 15–16 | 16–17 |
| Random           | CX 515.747 | 412.598 | 471.558 | 463.745 | 365.845 | 515.747 |
|                  | CP  | 89% | 100% | 80% | 91% | 90% | 71% |
| Moving Average   | CX 378.898 | 382.369 | 345.584 | 347.651 | 338.853 | 382.369 |
|                  | CP  | 95% | 99% | 100% | 90% | 91% | 89% |
| Exponential Smoothing | CX 471.558 | 497.798 | 256.737 | 237.038 | 208.317 | 452.591 |
|                  | CP  | 100% | 93% | 97% | 57% | 52% | 46% |

### Table 17. Cumulative slope index values for three models (2016.08–09) and base model (2016.10).

| Model Group      | Cumulative Slope Index | Maximum Load Time Zone | CX<sub>max</sub> |
|------------------|------------------------|------------------------|------------------|
|                  | 10–11  | 11–12  | 13–14 | 14–15 | 15–16 | 16–17 |
| Random           | CX 515.747 | 412.598 | 471.558 | 463.745 | 365.845 | 515.747 |
|                  | CP  | 89% | 100% | 80% | 91% | 90% | 71% |
| Moving Average   | CX 378.898 | 382.369 | 345.584 | 347.651 | 338.853 | 382.369 |
|                  | CP  | 95% | 99% | 100% | 90% | 91% | 89% |
| Exponential Smoothing | CX 471.558 | 497.798 | 256.737 | 237.038 | 208.317 | 452.591 |
|                  | CP  | 100% | 93% | 97% | 57% | 52% | 46% |

### Table 18. Cumulative slope index values for three models (2016.09–10) and base model (2016.11).

| Model Group      | Cumulative Slope Index | Maximum Load Time Zone | CX<sub>max</sub> |
|------------------|------------------------|------------------------|------------------|
|                  | 10–11  | 11–12  | 13–14 | 14–15 | 15–16 | 16–17 |
| Random           | CX 515.747 | 412.598 | 471.558 | 463.745 | 365.845 | 515.747 |
|                  | CP  | 89% | 100% | 80% | 91% | 90% | 71% |
| Moving Average   | CX 378.898 | 382.369 | 345.584 | 347.651 | 338.853 | 382.369 |
|                  | CP  | 95% | 99% | 100% | 90% | 91% | 89% |
| Exponential Smoothing | CX 471.558 | 497.798 | 256.737 | 237.038 | 208.317 | 452.591 |
|                  | CP  | 100% | 93% | 97% | 57% | 52% | 46% |
As shown in Tables 8–19, some values of CP parameter of cumulative slope index show that the predicted result of the method is analogous to the result of the representative electrical load or indicate that the method forecasts the maximum electrical load duration (cumulative slope index $\geq 80\%$) exactly while the representative electrical load was within the maximum electrical load duration.

In contrast, other values of CP parameter of cumulative slope index show that the method forecasts the opposite result compared to the representative electrical load.

Through analysis of the results in Tables 8–19, the cumulative slope index values for the random sampling technique show 86% rates analogous to the representative electrical pattern among 72 maximum load time zones. The moving average shows 69% and the exponential smoothing model shows 67%.

These results show that the proposed algorithm using random sampling has a more analogous representative electrical pattern for obtaining probabilistically obvious affinities with the patterns for other dates (correlation coefficient $\geq 0.8$) than do the other models (moving average and exponential smoothing).

According to the cost saving effects from quantitative evaluation, we performed an evaluation using the KEPCO electricity bill. Table 20 presents one quantitative evaluation and shows the evaluation result for 2016.04 for three models (random sampling, moving average, and exponential smoothing).

In the same way, the fee for the total duration (from 2016.01 to 2016.12: 12 months), can also be calculated. The evaluation results of 12-day fees and 240-day fees (weekdays over 12 months, excluding holidays) are shown in Table 21. As shown in Table 21, it was found that the proposed algorithm using random sampling could provide electricity cost savings of 0.62–2.28% greater than with the conventional electricity prediction techniques such as moving average and exponential smoothing for the total duration. Cost savings can be specified as the cost difference between moving average and random sampling, or between exponential smoothing and random sampling. Cost saving effects can be presented as cost saving percentages based on random sampling. When calculating the total
electricity bill, we did not include basic charges and only considered the unit price in various time zones. This is because we only wanted to discuss the quantitative performance and efficiency of the proposed algorithm using random sampling.

Table 20. Example for quantitative evaluation (2016.04) of three models.

| Model Group | Cumulative Slope Index | Maximum Load Time Zone |
|-------------|------------------------|------------------------|
|             | 10–11  | 11–12  | 13–14  | 14–15  | 15–16  | 16–17  |
| Random      | 92%    | 85%    | 90%    | 83%    | 70%    | 63%    |
| Moving Average | 86%    | 77%    | 79%    | 75%    | 75%    | 66%    |
| Exponential Smoothing | 80%    | 79%    | 77%    | 75%    | 75%    | 66%    |
| Representative Electrical Load | 95%    | 87%    | 90%    | 87%    | 100%   | 86%    |

Table 21. Cost savings and cost saving effects for quantitative evaluation of three models.

| Fee                  | Model                  | Costs(KRW) | Costs(USD) |
|----------------------|------------------------|------------|------------|
| Total fee (12 days)  | Random Sampling (RS)   | 19,544,636 | 18,477     |
|                      | Moving Average (MA)    | 19,666,534 | 18,562     |
|                      | Exponential Smoothing  | 19,990,568 | 18,867     |
| Total fee (240 days) | Random Sampling (RS)   | 390,892,717 | 368,940   |
|                      | Moving Average (MA)    | 393,330,672 | 371,241   |
|                      | Exponential Smoothing  | 399,811,369 | 377,358   |
| Cost savings between MA and RS (240 days) | 2,437,955 | 2,301 |
| Cost savings between ES and RS (240 days) | 8,918,651 | 8,417 |

Cost Saving Effects

| Cost savings (%) based on RS | Compared with MA | 0.62% |
|                             | Compared with ES  | 2.28% |

5. Conclusions

In this paper, we propose an algorithm for forecasting maximum electrical load duration using random sampling and a cumulative slope index. Before moving toward the long-term and ultimate solution such as machine learning, we propose a simple and efficient method to forecast electricity usage patterns and the duration of maximum electrical load using a small data set. The algorithm can forecast electricity usage patterns and maximum electrical load duration for a building with a constant lifestyle and electricity use pattern on a weekly basis utilizing a random sampling technique.

In contrast to sequential correlation analysis, for which we might examine as many data as possible to determine the pattern of electrical usage, the proposed algorithm does repetitive correlation analysis with data extracted by random sampling. With the proposed approach, this paper offers the potential for electrical use patterns to be obtained more quickly than with sequential correlation analysis, in terms of stochastic method.

To verify the effectiveness of the algorithm, we used electricity data (2015.11–2016.12) for the headquarters building of KEPCO KDN located in Naju City, South Korea, to forecast the electrical use pattern. We proposed an algorithm to provide an alternative approach to forecasting the electrical
To represent the usage data each month, employing random sampling and repetitive correlation analysis. To evaluate the performance of the proposed algorithm, we forecast the duration of maximum electrical load with each model (random sampling, moving average, and exponential smoothing) and the quantitative performance of the algorithm was evaluated using KEPCO electricity bills, ESS discharging condition, and cumulative slope index. As a result, it was found that the proposed algorithm can provide electricity cost savings of 0.62–2.28% greater than with other, more conventional electricity prediction techniques such as moving average and exponential smoothing. In near-future research, it is expected that this algorithm will become applicable to electrical applications involving big data, to deal with real-time data processing. It should also offer an alternative method for exploring the use of ESS systems for incorporation into various complex big data systems.

Author Contributions: J.K. conceived and designed the proposed algorithm using random sampling and cumulative slope index and wrote the paper. H.H. performed various experiments and analyzed the results with other methods. K.-I.K. supervised this research and reviewed the manuscript.

Funding: This research was funded by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry and Energy (MOTIE) of the Republic of Korea grant number [20168530050080].

Acknowledgments: This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry and Energy (MOTIE) of the Republic of Korea (No. 20168530050080).

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

References
1. Razmara, M.; Bharati, G.; Hanover, D.; Shahbakhti, M.; Paudyal, S.; Robinett, R., III. Building-to-grid predictive power flow control for demand response and demand flexibility programs. *Appl. Energy* 2017, 203, 128–141. [CrossRef]
2. Eissa, M. Real time incentive demand response program in smart grid with “i-Energy” management system with different resources. *Appl. Energy* 2018, 212, 607–621. [CrossRef]
3. Yu, M.; Hong, S. Incentive-based demand response considering hierarchical electricity market: a stackelberg game approach. *Appl. Energy* 2017, 203, 267–279. [CrossRef]
4. Motalleb, M.; Ghorbani, R. Non-cooperative game-theoretic model of demand response aggregator competition for selling stored energy in storage devices. *Appl. Energy* 2017, 202, 581–596. [CrossRef]
5. Hor, C.; Watson, S.; Majithia, S. Analyzing the impact of weather variables on monthly electricity demand. *IEEE Trans. Power Syst.* 2005, 20, 2078–2085. [CrossRef]
6. Amjady, N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Trans. Syst.* 2001, 16, 798–805. [CrossRef]
7. Taylor, J.; Buizza, R. Using weather ensemble predictions in electricity demand forecasting. *Int. J. Forecast.* 2003, 19, 57–70. [CrossRef]
8. Ramanathan, R.; Engle, R.; Granger, C.; Vahid-Araghi, F.; Brace, C. Short-run forecasts of electricity loads and peaks. *Int. J. Forecast.* 1997, 13, 161–174. [CrossRef]
9. Oliveira, E.; Oliveira, F. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy* 2018, 144, 776–788. [CrossRef]
10. Dong, Z.; Yang, D.; Reindl, T.; Walsh, W. Short-term solar irradiance forecasting using exponential smoothing state space model. *Energy* 2013, 55, 1104–1113. [CrossRef]
11. Taylor, J. Triple seasonal methods for short-term electricity demand forecasting. *Eur. J. Oper. Res.* 2010, 204, 139–152. [CrossRef]
12. Taylor, J. Short-term electricity demand forecasting using double seasonal exponential smoothing. *J. Oper. Res. Soc.* 2003, 54, 799–805. [CrossRef]
13. Amjady, N. Short-term bus load forecasting of power systems by a new hybrid method. *IEEE Trans. Power Syst.* 2007, 22, 333–341. [CrossRef]
14. Silva, D.; Yu, X.; Alahakoon, D.; Holmes, G. A data mining framework for electricity consumption analysis from meter data. *IEEE Trans. Ind. Inform.* 2011, 7, 399–407. [CrossRef]
15. Zhou, H.; Lin, B.; Qi, J.; Zheng, L.; Zhang, Z. Analysis of correlation between actual heating energy consumption and building physics, heating system, and room position using data mining approach. *Energy Build.* **2018**, *166*, 73–82. [CrossRef]

16. Suomalainen, K.; Pritchard, G.; Sharp, B.; Yuan, Z.; Zakeri, G. Correlation analysis on wind and hydro resources with electricity demand and prices in New Zealand. *Appl. Energy* **2015**, *137*, 445–462. [CrossRef]

17. Kim, A.; Lee, S.; Park, J.; Kang, H.; Seong, P. Correlation analysis between team communication characteristics and frequency of inappropriate communications. *Ann. Nucl. Energy* **2013**, *58*, 80–89. [CrossRef]

18. Borges, C.; Dias, J. A model to represent correlated time series in reliability evaluation by non-sequential Monte Carlo simulation. *IEEE Trans. Power Syst.* **2017**, *32*, 1511–1519. [CrossRef]

19. Wangdee, W.; Billinton, R. Bulk electric system well-being analysis using sequential Monte Carlo simulation. *IEEE Trans. Power Syst.* **2006**, *21*, 188–193. [CrossRef]

20. Baležentis, T.; Streimikiene, D. Multi-criteria ranking of energy generation scenarios with Monte Carlo simulation. *Appl. Energy* **2017**, *185*, 862–871. [CrossRef]

21. Zio, E.; Delfanti, M.; Giorgi, L.; Olivieri, V.; Sansavini, G. Monte Carlo simulation-based probabilistic assessment of DG penetration in medium voltage distribution networks. *Electr. Power Energy Syst.* **2015**, *64*, 852–860. [CrossRef]