# Time Series Outlier Detection for Short-Term Electricity Load Demand Forecasting

**Chawalit Jeenanunta[1], K. Darshana Abeyrathna[2], M. H. M. R. Shyamali Dilhani[3], Su Wutyi Hnin[1] and Pyae Pyae Phyo[1]**

[1]School of Management Technology, Sirindhorn International Institute of Technology, Thammasat University, Pathumthani 12000, Thailand
[2]Information and Communication Technology, University of Agder, Norway
[3]Department of Interdisciplinary Studies, Faculty of Engineering, University of Ruhuna, Sri Lanka
E-mail: chawalit@siit.tu.ac.th, darshana.abeyrathna@uia.no, rasidilhani@gmail.com, d6022300203@g.siit.tu.ac.th, pyaepyaephyo1979@gmail.com

*Abstract*—**Forecasting of working-days' electricity demand is vital for short-term planning. However, demand variations due to outliers can reduce the accuracy of forecasts. Therefore, a time series data cleaning technique is proposed to remove these disturbances of electricity data. First, holidays' and bridging holidays' data are replaced by Moving Average. The k-sliding window filtering band is proposed to detect the time series outliers and replace by forecasted regular load demand using Moving Average. Data from the Electricity Generating Authority of Thailand (EGAT) and a Neural Network (NN) model with six inputs and one output are used to demonstrate the performance of time window data cleaning process. The sample dataset contains data from 1stMay 2012 to 31stMay 2013 where May 2013 is used for testing. The Time-Window based data cleaning technique increases the performance of forecasting outcomes by 11.60% for non-holidays. Results from the proposed technique are compared with the results from the robust version of locally weighted smoothing (r-LOESS) and identified that the proposed technique is superior for taking results for non-holidays.**

*Index Terms*—**Data Preprocessing, Data Cleaning, Outliers Detection, Filtering Band, Short-Term Load Forecasting and Neural Network**

## I. Introduction

Load forecasting is an indispensable section on designing, planning, and operating of electric utilities. It is generally classified according to the time interval into three main classes; Short-Term Load Forecasting (STLF) ranging from one hour to one week, Middle-Term Load Forecasting (MTLF) ranging from one month up to one year, and Long-Term Load Forecasting (LTLF) ranging for more than one year. The conventional time series models or computational intelligence-based models are used for forecasting under the above three classes. Traditional time series models such as regression analysis [1-4], moving average [5], and stochastic time series [6-8] have been applied by several researchers. Artificial intelligence [9-15], deep learning [16], machine learning [17-19], fuzzy time series [20, 21], and expert systems [22] have been appliedunder intelligence based models [23].

The load data consist of repeated data patternsdue to daily, weekly, and seasonally variations [24, 25]. However, identifying of these patterns is a quite complex task because load data are affected by different other factors such as weather, temperature, and unusual consumption patterns. The quality of data is also complex the forecasting process. Therefore, data preprocessing plays an important role in electricity forecasting. It consists noise removal and data cleansing.

A large amount of data can consist outliers due to randomness or noises. All these unpredictable patterns and undetectable outliers can reduce the forecasting accuracy. Proper detection of outliers and replacing them with clean data help to improve the forecasting accuracy. Therefore, this research proposes a data preprocessing step before using itin the forecasting stage. There are two types of outlier detection methods called univariateand multivariate methods [26]. Examplesfor univariate outlier detections are Single-step Procedures, Sequential Procedures, Inward and Outward Procedures, Univariate Robust Measures, Statistical Process Control (SPC). Example for multivariate methods is Statistical Methods, Multivariate Robust Measures, Data-Mining Methods, and Preprocessing Procedures. However, the Time-Window based model proposed in this research is novel and contrastive to the above methods.

In the second section of this paper, related works on detecting outliers and replacing outliers are reviewed. The proposed data cleaning process are discussed in the third section Then, Artificial Neural

Networks for forecasting are discussed in the fourth section. In the methodology section, the design of experiments to evaluate the performances of data cleaning process is explained. Results of the research are discussed in the next section. Finally, conclusions are stated based on the obtained results for then tire research.

## II. Literature Review on Data Cleansing Techniques

### A Outlier Detection

The outliers can be identified as the data that deviate so much from the other data as they have generated by a different mechanism [27]. In other words, outliers are the data which exhibit an inconsistent behavior compared to the remaining data set [28]. Outliers can be categorized into two groups: outliers in the x-axis and outliers in the y-axis. However, in time series analysis, outliers in y-axis are more important which can be re-categorized into another three groups called random, non-random, and gross errors [29]. Thus, the identification and treatment of outliers are important for the time series analysis before identifying the patterns in them.

Electric energy consumption is recorded for daily operations such as system analysis, visualization, reliability performance, energy saving, and system planning. However, it cannot be avoided that historical load data consist missing values and corrupted values due to the random failures in metering and transferring processes. Conventional outlier detection methods are based on the assumptions such as raw data is identically and independently distributed. Usually, statistical distribution methods are not used to detect outliers because it has bulk data. However, numerous outlier detection methods have been used by researchers for different applications such as clinical trials, weather prediction, and electricity forecasting. Especially in forecasting, historical data with outliers cause for low forecasting accuracy. Proper outlier detection methods can remove those outliers and increase the forecasting accuracy. Therefore, following powerful data cleaning processes is an important phase in forecasting [30].

Liu, Shah [31] proposed an on-line outlier-resistant model combining with a modified Kalman filterto detect and clean outliers. Thesugges ted filtering technique is simple and reliable since it has a high break down points and one or two parameters.Trueck, Weron [32] discussed different outlier detection techniques such as low-pass filtering, percentage thresholds, and fixed price thresholds to detect the outliers in the electricity price data. Hadi [33] introduced a model comprising several steps to identify multiple outliers in multivariate data. The model is started with arranging data into the ascending order. In addition to that, Basic Subset of Full Rank, Basic Subset Not of Full Rank, and Increase Size of Basic Subset are the other steps to identify outliers. Hadi [34] proposed a modified model changing the parameters in his previous model to get better re sultsin detecting multiple outliers in multivariate data.

Janssens et al., [35] proposed outlier detection methods to utilize Machine Learning (ML) and Knowledge Discovery in Databases (KDD). They compared the Support Vector Machines Data Description (SVDD), Parzen Windows (PW), and K-Nearest Neighbors (KNN) methods for the field of ML and the Local Correlation Integral (LCI) and the Local Outlier Factor (LOF) methods for KDD. Another technique is Replicator Neural Networks (RNNs) algorithm which has been introduced to detect the outliers of both small and large datasets [36, 37]. During the training process, common patterns have a higher impact on adjusting the weights while the patterns representing outliers have a less impact. There fore, during the testing phase, the data with higher errors are identified as outliers. Three statistical outlier detection techniques namely, cluster outliers, radial outliers, and scattered outliers are set as a benchmark for the proposed model.

Chen, Li [38] had used B-Spline smoothing based techniques to clean the corrupted and missing values in the data from British Columbia Transmission Corporation (BCTC) while [39] used B-spline approach to clean the temperature and historical load data as a preprocessing step for short-term (24-hour) load prediction in British Columbia. Experimental resultsindicate that the overall performance of the load prediction has improved after the data cleansing process.Tang, Wu [40] discusseda new approach for outlier detection of load data under three different casesnamely, Outlier Detection for Normal Distributed Data, Outlier Detection for Gamma Distributed Data, and Outlier Detection for Small-Size Portrait Data. Outlier detection using small size portrait data-based approach significantly improves the outliers' detection process compared with other approaches.do Nascimento, Oening [41] summarized diverse outliers' detection and filling algorithms for load time series data in the metering centers of smart grids called Extreme Studentized Deviate (ESD), Generalized Extreme Studentized Deviate (GESD), Z-Score, Test Box Plot, Thompson, Adjusted Box plot, and Exponential Smoothing. Considering the problem as the presence of null values in the readings, they propose a method using the modified Z-Score technique. The proposed method detects outliers at two stages called Pre-Detection and Post-Detection.

Statistical inference methods for power monitoring tasks against the outlier effects due to faulty readings and malicious attacks were discussed byMateos and Giannakis [42]. They developed a novel load cleansing and imputation scheme leveraging the low intrinsic-dimensionality of spatiotemporal load profiles. The concept of a robust estimator based on

Principal Components Pursuit (PCP) is usedin the technique. Finally, Hodge's review [43] identifies causes for outliers such as changes in system behavior, instrument errors, fraudulent behavior, and mechanical faults. He summarized outlier detection techniques and brought the idea that principled and systematic techniques are highly used in the recent research.

*B. Outlier Estimation*

A research had been conducted to replace the outliers and missing values of time series related to the sludge wastewater treatment plant in Edinburgh, U.K[44]. Due to the limited amount of data, missing values and outliers cannot be discarded. Therefore, they suggest the Kohonen Self-Organizing Map (KSOM: one of Artificial Neural Network algorithms) technique and unsupervised neural networks to forecast the missing values. They identify the advantages of using KSOM as higher accuracy, computationally efficient, and the simplicity. However, usage of back propagation artificial neural network in replacing outliers decreases the performance of the model when the number of output variables is particularly high [45] or output variables of the ANN are not highly correlated [46]. Simple linear regression methods to estimate the missing values are discussed by MacDonald and Zucchini [47] and Harvey [48]. They identified that the results are getting better when the series has less missing data points while they use their method to check the water quality parameters.

Apart from these specific outlier replacing methods, researchers have used some basic forecasting techniques to replace the missing values and outliers: Moving Average [26], Exponential Smoothing [49], Regression methods [50]. Some of them had used some advanced techniques such as Neural Networks, Machine Learning for the same purpose. Selected outlier replacing technique used in this research is discussed in the methodology.

## III. Data Cleaning Process

In this section, we discuss the proposed Time-Window based data cleaning process for load and temperature data. The load data obtained from Electricity Generating Authority of Thailand (EGAT) consist of abnormal patternsdue to holidays, outliers, or missing values. However, temperature data has only outliers and missing values as they do not change with holidays. Therefore, these irregular rpatterns of both load and temperature data have to beidentified and replaced by estimated data. A sample data set is selected from 1st of April 2012 to 31st of May 2013 from Bangkok and Metropolitan regions.

*A. Detecting and Replacing Holidays*

Holidays are one of the major reasons to fluctuate the regular electricity consumption and it is the easiest type to identify. Connor[51]showed that the proposed Neural Network based outlier detection technique can detect holidays as outliers during the research. Since almost all the factories and industries do not continue their works on holidays, there is an obvious electricity demand depression on holidays. Including the holiday demand values in the training data can reduce the forecasting results' accuracy. Therefore, as the first step of the data cleaning process, holidays have to be identified and replaced by forecasted or estimated demand. Simply the calendar holidays in the selected period are identified and replaced by the Weighted Moving Average method as given in Eq. (1).

$$L_t(d) = [w_1 * L_t(d-7) + w_2 * L_t(d-14)] \tag{1}$$

Where, $L_t(d)$ is the electricity load of day $d$ at time period $t$. The regular demand from the previous two weeks is selected as they have the same pattern throughout the day. $w_1$ and $w_2$ are the adjustable weight values. For the most recent week has the most relevant information, $w_1$ and $w_2$ are fixed at 0.7 and 0.3.

*B. Detecting and Replacing bridging holidays*

When there is a non-holiday between two holidays or when there is a non-holiday between a holiday and a weekend, we call it as a bridging holiday. Normally, people tend to take leaves on bridging holidays to extend their holiday-period. As a result, unexpected demand decrements can be seen compared to the regular non-holiday demand. Therefore, we identify and replace them using the same method used to replace the calendar holidays as the second step of the data cleaning process.

*C. Detecting and Replacing Outliers*

Other than the load values in holidays and bridging holidays, some load values show irregular patterns (e.g. due to measurement errors). Finding such abnormal load values as outliers and replacing them by estimated load will improve the forecasting performance. Therefore, the third preprocessing step is to detect and replace these outliers.

*1. Time Series Outlier Detection*

A time series is a set of observations which are collected between equal time intervals. The electricity data use in this research has been gathered for every 30-minute. Therefore, the time series of the load consumption can be arranged in many ways by dividing them into different time intervals Fig. I. shows three possible ways of arranging electricity consumption data. The first graph (a) is drawn using the data at 11.00 a.m. of all Fridays from 1st of May, 2012 to

31st of May 2013. The second graph (b) has data at 11.00 a.m. of all the days from 1st of May, 2012 to 20th September 2012 while the third graph (c) consists with the data of all the time periods of all the days from 1st of May, 2012 to 7th of May, 2012. The standard deviations of the given series are 427.82, 975.88, and 1197.50, respectively. This reveals that the 2nd and the 3rd series have higher demand variations compared to the 1st series. Therefore, detecting outliers in the second and third series is not easy as they have higher standard deviations which lead to weakening the outlier detection process. Outliers in the first series are easier to recognize as it has a stable series.

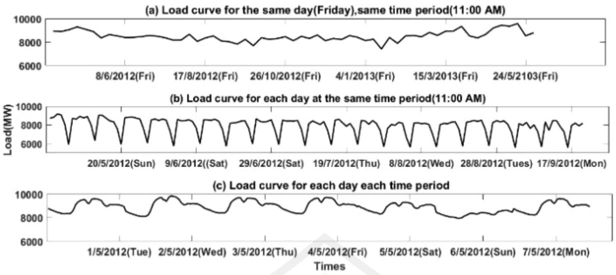


Fig. I. Load demand variation of each time period arranging

A filtering band is used to classify the outliers in the time series where the data that are outside the filtering band areclassified as outliers. The previous works on the filtering bands were constructed by using all data in the time series (all *d* with all *t*) as the data arrangement given in the series (c) [38, 45, 52]. However, to minimize the error of misclassifying outliers as non-outliers, the data is arranged as given in the series (a) of the Fig. I. There fore, the time interval of the time series is one week and $7 \times 48$ individual Time-Window vectors are identified. These data from different Time-Windows with one week time intervals are put in the vectorsas given in Eq. (2).

The set of days in each day of week can be defined as $\{d', d'-7, \ldots, d'-7m\}$ and $d'$ represents the last 7 days in the selected data set.

$$V_t(d') = [L_t(d'), L_t(d'-7), L_t(d'-14), \ldots L_t(d'-7 \times m)] \quad (2)$$

Where, $V_t(d')$ is the vector for Time-Window of day, $d'$, and time period, *t*, for*m* weeks. Where *m* is equal to 60 for the vectors with $d'$ equal to Saturday as thereare 60 Saturdays in the selected sample data set and *m* is equal to 61 for all the other vectors.

*k* period filtering bands for each Time-Window vector is created. The *k* period filtering band of time period *t* on day $d'$, $B_t(d')$,is given in Eq. (3). It is constructed with the $k^{\text{th}}$period moving average and the standard deviation of the vector or, $SD(V_t(d'))$, which is calculated for the vector $V_t(d')$.

$$B_t(d') = \frac{[\sum_{i=1}^{k} L_t(d'-7 \times i)]}{k} \pm N \times SD(V_t(d'));$$
$$t = 1, \ldots, 48, \forall d' \quad (3)$$

The width of thek-sliding window filtering band varies with the size of *N*. Small *N* detects a large number of outliers which leads to detect non-outliers as outliers. At the same time, large *N* detects a small number of outliers which leads to miss the actual outliers. The optimal *N* in this research is used as 1.6. The moving average is calculated by previous recent four weeks ($k = 4$).

*2. Replacing Outliers*

Outlier replacement is the most important part before continuing the forecasting phase. In this paper, we discuss two data replacement techniques; Moving Average (MA) and Interpolation.

*a. Moving Average (MA)*

Moving average is a simple time series forecasting method where it is used to estimate the regular demand of the identified outliers in the previous section. The arithmetic moving average is selected due to its simplicity as given in Eq. (4).The average of the load of *n* previous weeks of day, *d*, at the time period, *t*, is used to replace the identified outliers.

$$L_t^p(d) = \frac{[\sum_{i=1}^{n} L_t(d-7 \times i)]}{n} \quad (4)$$

Where,

$L_t^p(d)$= Forecastedregular load of day *d* at time period *t*

*b. Interpolation*

Another simple method is linear interpolation which is used to find the values between the data points. The points are connected in a simple manner joining by straight line segments. Each segment bounded by two points can be interpolated independently. The purpose of the interpolation is to replace a set of data points with a function given analytically.

$$L_t^p(d) = L_t(d-7) + \frac{m_2 - m_1}{(m_3 - m_1)}(L_t(d+7) - L_t(d-7)) \quad (5)$$

Where, $L_t(d-7)$ = Previous week same day load at time period $t$

$L_t(d+7)$ = Next week same day load at time period $t$

$m$ = The time axis positions of load values where $m_1$ is the position of $L_t(d\text{-}7)$, $m_2$ is the position of $L_t(d)$, and $m_3$ is the position of $L_t(d\text{+}7)$.

The vector of Wednesday at 1.30 p.m. is drawn in Fig II. The interpolation just connects the adjacent points and completes it by a straight line while there is an acceptable variation when outliers are replaced by the moving average.

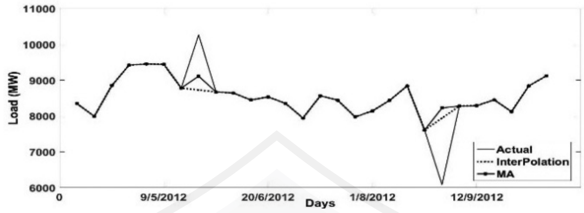


Fig. II. Replacing outliers by MA and Interpolation

Therefore, moving average is used as the outlier replacement method in all the other the future experiments of this research. One of the strong points of Time-Window based data cleaning technique is, even if we miss holidays or bridging holidays and leave them without replacing, those days can be detected by the proposedk-sliding window filtering band.

*3. r-LOESS*

Locally weighted smoothing or the LOESS is a special case of Linear Regression. Instead of finding the relationship for the entire dataset directly, LOESS considers the surrounded points within a span to make the curve smoother than Linear Regression does. Weight values for the points within the span are calculated using a regression weight function. These weight values along with the magnitude of the surrounded data points define the value to be smoothed while there is no influence by the points outside the span [53, 54].

Outliers in the data set can distort the shape of the LOESS curve. Therefore, the robust smoothing procedure is applied to LOESS (r-LOESS) with the purpose of eliminating the effect of outliers on LOESS. The robust smoothing helps to put zero weights on the outliers and then there is no influence by outliers on the data to be smoothed.[55] The robust procedure calculates the Mean Absolute Deviation (MAD) of the data within the span. All the data outside the $6 \times MAD$ is identified as outliers and put zero weights on them. Finally, the weighted regression analysis is performed within the span and this process is repeated 5 times before getting the final results.

After replacing holidays and bridging holidays using the Eq. (1), all the data is taken to a single vector as given in series (c) of Fig I. This vector is fed into the MATLAB software for using it with the "rloess" pre-defined function.

*4. Detecting and Replacing Outliers of Temperature Data*

Considering the smooth variations in temperature curves throughout the months, weeks, and days, data are not separated into separate vector sand consider only one Time-Window. Since holidays do not make changes on temperatures, raw temperature data is directly checked with the filtering band as given in Eq. (6).

$$BT_t(d) = \frac{[\sum_{i=1}^{k} T_{t-i}(d)]}{k} \pm N \times SD(V(T)); \quad t = 1, \ldots, 48 \quad (6)$$

The filtering band of time period $t$ on day $d$, $BT_t(d)$ for temperature data is created using the above equation. $T_t(d)$ is the temperature value at time $t$ on day $d$ and $V(T)$ is the vector for the only Time-Window that contain all the temperature data. Therefore, the time interval for the time series for the temperature $V(T)$ is 30-minutes. When $k$ = 4, the filtering band $BT_t(d)$ is defined by the four period moving average and the standard deviation of the vector $V(T)$. These identified outliers in temperature data are replaced by moving average method as given below.

$$T_t^p(d) = \frac{[\sum_{i=1}^{n} T_t(d-i)]}{n} \quad (7)$$

Where $T_t^p(d)$ is the estimated temperature value for time period $t$ on day $d$. This is calculated with $n = 3$, as it takes the average of the same time period from 3 recent days.

The complete process from raw data to evaluating forecasting performance with the Time-Window based data cleaning process is illustrated by the following steps and evaluation process with its results are discussed in the following section.

1. Raw Data
   a. Load values for each 30-minute from March 2009 to December 2013 are gathered by EGAT
   b. A sample from April 2012 to May 2013 is selected in this research
2. Replacing Holidays
   a. Calendar holidays from April 2012 to April 2013 are identified
   b. Replace each time period of holidays by Eq. (1)
3. Replacing Bridging Holidays
   a. Days between two holidays or holidays and weekends are identified as bridging holidays
   b. Replace each time period of bridging holidays by Eq. (1)
4. Detecting outliers
   a. Average of adjacent periods and Standard Deviation of data series are calculated to construct the equation (3) and (6) for each Time-Window
   b. Eq. (3) and (6) are used to identify the outliers of load and temperature, respectively
5. Replacing outliers
   a. Detected outliers are replaced by moving average (Eq. (4) and (7) for load and temperature, respectively)
6. Training the forecasting models
   a. Data are separated into each time interval (48)
   b. Data from May 2012 to April 2013 are selected as the training data
   c. To forecast one day, 48 training sets are required with separate models for each time interval
   d. Neural Network sare used in the forecasting phase
7. Test with unseen data
   a. May 2013 is selected as the testing month
   b. Each day, each time period is forecasted with the suggested method and calculate MAPEs for actual and forecasted loads

## IV. Forecasting Technique

This research is conducted to forecast the next day electricity demand consisting of 48 periods. As given in the Eq. (8), previous day load, previous week load, previous day temperature, and the same day forecasted temperature for the same time interval are used to forecast the next day electricity demand. In addition to that, we use the *Day of Week* and *Month of Year* for recognizing the weekly and monthly demand variations of the data.

$$\begin{aligned} F_t(d) = {} & a_1 L_t(d-1) + a_2 L_t(d-7) \\ & + a_3 T_t(d) + a_4 T_t(d-1) \\ & + a_5 \mathit{Day\ of\ Week} \\ & + a_6 \mathit{Month\ of\ Year} \end{aligned} \tag{8}$$

$F_t(d)$ = Forecasted load on day $d$ at period $t$
$L_t(d)$ = Actual load on day $d$ at period $t$
$T_t(d)$ = Temperature on day $d$ at period $t$
$(a_i)$ = Coefficients which are represent the weight values of the NN, where $i$ = 1,2,3,4,5,6
*Day of Week* = 1-Sunday, 2-Monday, …, 7-Saturday
*Month of Year* = 1-January, 2-February, …, 12-December

Since Neural Networks (NNs) have the ability of learning and recognizing non-linear patterns of complex data sets, the NN is used in this research to demonstrate the improvement of data preprocessing. An example of a simple NN structure is shown in Fig.III. Based on the data arrangement given in Eq. (8), the Neural Network, $NN_t^d$ is trained separately for forecasting the load values of period $t$ on day $d$. For each $NN_t^d$, there aresix input nodes and one output node.

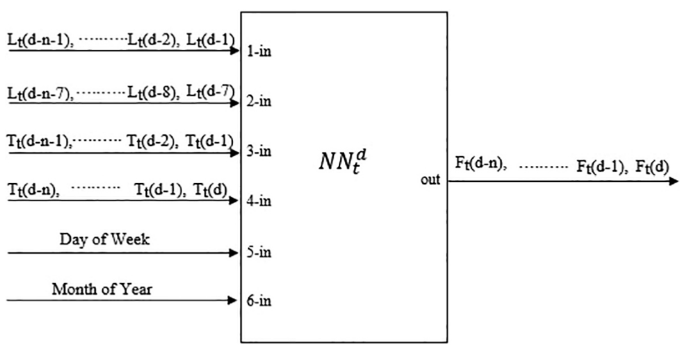


Fig. III. Artificial Neural Network input and output structure

## V. Methodology

Data from Electricity Generating Authority of Thailand (EGAT) is used to test the Time-Windowbased data cleaning technique. A sample dataset is selected for training and testing the NN where it contains Bangkok and Metropolitan region's data ranging from 1st of May, 2012 to 31st of May, 2013 for every 30-minute. One-year data is used to trainthe selected forecasting model. The model performance is tested on May data as it is one of the hardest months to forecast due to several reasons. The the temperature in Thailand is at the peak in May. Moreover, forecast performance of May always gets higher errors in forecasting as April data has a lot of holiday sand it is included in the training dataset. Therefore, May is selected as the testing month to demonstrate the performance of the Time-Window based technique.

The example data arrangement given in Table I is for the testing data of May 1, 2013. It includes six inputs of data for a given forecasting load. The same structure can be used to forecast the other days in the testing month. There are different 358 training datasets for each of the testing day. The same NN parameters and the structure are used with both raw and preprocessed data sets. The number of hidden layers is equal to 1, the number of neurons is equal to 7, and the number of epochs is equal to 1000.

TABLE I
An example of training and testing dataset

| | No. | Input-1 [Lt(d-1)] | Input-2 [Lt(d-7)] | Input-3 [Tt(d-1)] | Input-4 [Tt(d)] | Input-5 [DoW] | Input-6 [MoY] | Target Lt(d) |
|---|---|---|---|---|---|---|---|---|
| Training Data | 1 | 7/5/2012 (Wed) | 1/5/2012 (Thu) | 7/5/2012 (Wed) | 8/5/2012 (Thu) | 5 | 5 | 8/5/2012 (Thu) |
| | 2 | 8/5/2012 (Thu) | 2/5/2012 (Fri) | 8/5/2012 (Thu) | 9/5/2012 (Fri) | 6 | 5 | 9/5/2012 (Fri) |
| | 3 | 9/5/2012 (Fri) | 3/5/2012 (Sat) | 9/5/2012 (Fri) | 10/5/2012 (Sat) | 7 | 5 | 10/5/2012 (Sat) |
| | 4 | 10/5/2012 (Sat) | 4/5/2012 (Fri) | 10/5/2012 (Sat) | 11/5/2012 (Sun) | 1 | 5 | 11/5/2012 (Sun) |
| | | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
| | 358 | 29/4/2013 (Wed) | 23/4/2013 (Thu) | 29/4/2013 (Wed) | 30/4/2013 (Thu) | 5 | 4 | 30/4/2013 (Thu) |
| Testing-Data | No. | Input-1 [Lt(d-1)] | Input-2 [Lt(d-7)] | Input-3 [Tt(d-1)] | Input-4 [Tt(d)] | Input-5 [Dow] | Input-6 [Moy] | Output Ft(d) |
| | 1 | 30/4/2013 (Thu) | 24/4/2013 (Fri) | 30/4/2013 (Thu) | 1/5/2013 (Fri) | 6 | 5 | 1/5/2013 (Fri) |

The walk-forward testing routine is applied to test the performance of the 1- year training data set with 1 testing data to set. The data consist of the training and testing datasets. Each training data set consists of 358 pairs and each testing datasetconsists of 1 pair. Figure IV shows that for each experiment, the testing data slides one pair forward (1-31 testing dataset) and the ANN is trained with the new training data set that also slide forward to the future pair of data.

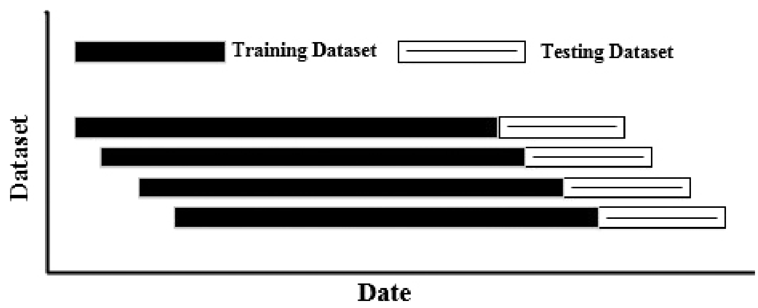


Fig. IV. The walk forward routine for dividing the dataset

In this research, Mean Absolute Percentage Error (MAPE) is used to evaluate the forecasting accuracy. MAPEs for all the testing days are calculated using the Equation (9).

$$MAPE_d = \left[ \left( \frac{1}{48} \right) \times \sum_{t=1}^{48} \left| \frac{e_t(d)}{L_t(d)} \right| \right] \times 100 \quad (9)$$

The MAPE can be calculated using the error, $e_t(d) = F_t(d) - L_t(d)$ where $F_t(d)$ is the forecasts obtained with regular load demand, $e_t^p(d) = F_t^p(d) - L_t(d)$ where $F_t^p(d)$ is the forecasts obtained with cleaned or replaced load with the proposed Time-Window based data cleaning technique, and $e_t^r(d) = F_t^r(d) - L_t(d)$ where $F_t^r(d)$ is the forecasts obtained with smoothed data with r-LOESS. $MAPE_d$, $MAPE_d^p$, and $MAPE_d^r$ for all the testing days ($d$ = 1,2,3….,31) in May, 2013 are compared in the results section.

## VI. Results

The data set for outlier detection and replacing has 426 days with a total of 20,448 demand values from April 1, 2012 to May 31, 2013. Demand values of April 2012 are also included in the data set as they required to replace the outliers in May 2012. There are 39 calendar holidays and 18 bridging holidays within the selected period. Therefore, 2736 of demand values are replaced by weighted moving average. Additionally, 1200 outliers are identified and replaced by two-period moving average.Since holidays and bridging holidays do not make any change in temperature values, raw temperature data are directly checked with the filtering band. These results, before moving them with NN are summarized in Table II.

TABLE II
Summary of outlier detecting and replacing

| | Total Days | Total Points | No. of Holidays | No. of Holiday points | No. of B. Holidays | No. of B. Holiday points | No. of Outliers |
|---|---|---|---|---|---|---|---|
| Load | 426 | 20,448 | 39 | 1,872 | 18 | 864 | 1200 |
| Temperature | 426 | 20,448 | | - | - | - | 1074 |

Figure V. shows the load curve before and after replacing holidays, bridging holidays, and outliers. December 8, 2012 is a Saturday and no outliers have been identified. A few outliers on Sunday, December 9, 2012 have been identified and replaced by the estimateddata. A higher number of outliers on Monday, December 10, 2012have been identified and replaced byestimateddata. Likewise, all the outliers are identified by the k-sliding window filtering band and replaced by the estimated data.

Figure VI shows the curve after replacing outliers with its k-sliding window filtering band. Since the outliers have already been replaced by two-period moving average, the curve is well within the k-sliding window filtering band. Even though the graph in Figure VI has been drawn with all data for just to be clear, all outliers are detected and replaced as explained in 4.3.1 and 4.3.2.

Considering the smooth variations in temperature curves throughout the months, weeks, and days, raw temperature values are directly checked with the filtering band. Figure VII shows the actual temperature curve before replacing the outliers and after replacing outliers with the filtering band. During the given period of the figure, few sudden reductions of temperatures have been identified and replaced by the moving average.

MAPE values before and after replacing outliers with the Time-Window based data cleaning technique and r-LOESSfor all the testing days in May, 2013 are summarized in Table III. $MAPE_d^p$ of almost all the non-holidays have been reduced. However, $MAPE_d^p$ of all the holidays have been increased by considerable amounts becausethe forecasting model is trained to forecast only the non-holidays' load. The minimum $MAPE_d$ is recorded on Sunday, May 12, 2013 (1.9381) and the maximum (21.7383) is on Wednesday, May 1, 2013 which is a holiday. The minimum $MAPE_d^p$ and $MAPE_d^r$ are recorded on Wednesday, May 29, 2013 (1.3801) and Sunday, May 12, 2013 (2.1669), respectively. All the maximum $MAPE_d$ (21.7383), $MAPE_d^p$ (26.0934), and $MAPE_d^r$ (19.2670) are recorded on Wednesday, May 1, 2013. $MAPE_d$ tend to be high on holidays and$MAPE_d^p$ and$MAPE_d^r$are even higher as seenin May 24, 2013 (Friday).

May 25, 2013 (Saturday)is not a holiday but still gives a higher $MAPE_d^p$and $MAPE_d^r$ even though the $MAPE_d$ is low. $MAPE_d^p$ and the $MAPE_d^r$ of May 25, 2013 are examples to show the relationship between the total load, peak load, and the $MAPE_d^p$ and $MAPE_d^r$. According to the Table III, the $MAPE_d^p$ and $MAPE_d^r$ are higher compared to $MAPE_d$ of a particular day if the total load and the peak load is low compared to the other days of May 2013. However, May 24, 2013

(Friday) is a holiday and this can be the reason to have a low electricity consumption on May 25, 2013 (Saturday) compared to the other Saturdays in the same month. This demand variation on May 25, 2013 (Saturday) compared to the other Saturdays is illustrated by Fig.VIII.

Figure IX shows how outlier replacement helps in forecasting on non-holidays. The best $MAPE_d^p$ is recorded on May 29, 2013 which is equal to 1.3801. Therefore, the forecasted outputs before and after replacing outliers have been plotted against the actual electricity consumption on Wednesday, May 29, 2013.

The forecasted curve which is after replacing outliers with the proposed Time-Window based technique behaves almost the same with the actual curve. However, forecasted curve using the raw data is significantly below the actual curve and shows the instability throughout the day.

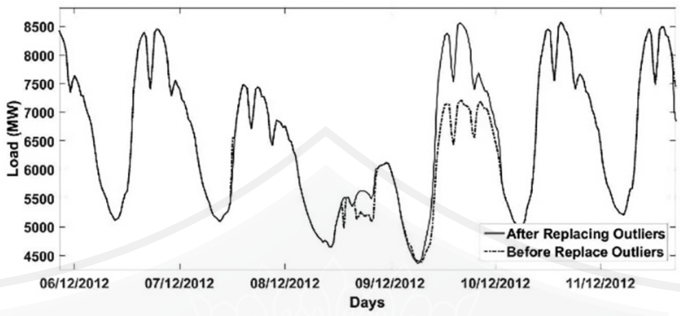


Fig. V. The load curve before and after replacing all outliers

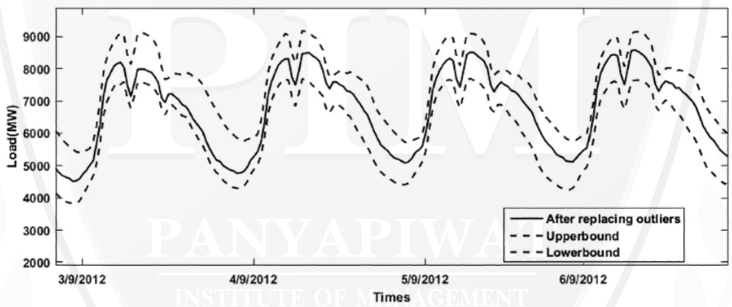


Fig.VI. The load curve after replacing outliers with the Time-Window based data cleaning technique

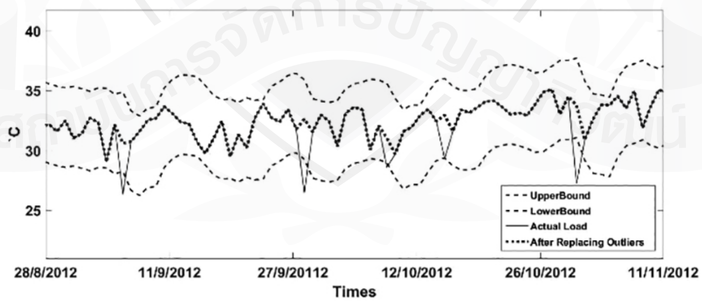


Fig.VII. Replaced outliers of the temperature data with the proposed filtering band

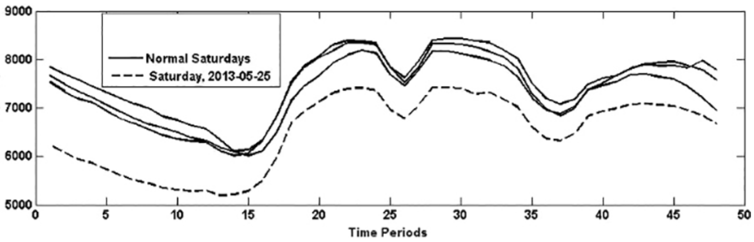


Fig. VIII. Demand depression on Saturday, 25.05.2013 compared to the other Saturdays

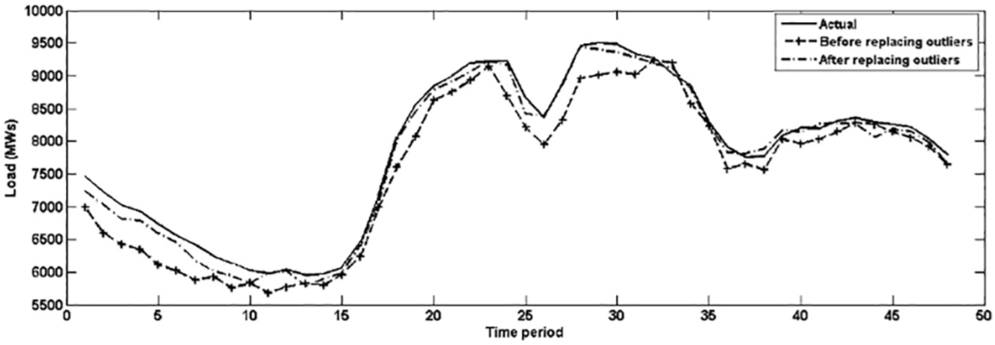


Fig. IX. Forecasted curves for a non-holiday (29/05/2013) before and after replacing outliers

## VII. Discussion

The results given in Table III are summarized in Table IV to get a clear idea about the effect of data cleaning on electricity forecasting accuracy. According to the Table IV, non-holidays' MAPE can be re-categorized into Monday's MAPE, Weekdays' (from Tuesday to Friday) MAPE, and weekends' MAPE. For the decrease of average non-holidays' MAPE with both Time-Window based and r-LOESS techniques there is a higher effect from average Mondays' MAPEs which are reduced from 4.5231 to 2.5830 and 4.5231 to 3.7532, respectively. The average MAPEs for other weekdays are reduced from 4.6810 to 4.1798 with the Time-Window based technique and reduced from 4.6810 to 4.1980 with r-LOESS. The average MAPEs for weekends is reduced from 4.2141 to 3.9430 when the data is cleaned with the Time-Window based technique and this value is reduced from 4.2141 to 3.5500 when data is cleaned with r-LOESS. Again, the holidays and outliers are disturbances in the training data when the NN forecasts the non-holidays' demand. Therefore, removing holidays and outliers from the training data set and replacing them by non-holidays help removing the disturbances in the training data. Therefore, this improves MAPEs for non-holidays.

However, the total average $MAPE_d^p$ is higher compared to the $MAPE_d$ and $MAPE_d^r$, but this does not affect to the research's objective. The increase of the total average $MAPE_d^p$ has a higher effect from $MAPE_d^p$ of holidays. Average holidays' MAPE is increased from 8.9902 to 11.8418, when the holidays and outliers are detected and replaced by the Time-Window based technique. This value is increased from 8.9902 to 9.7366 when the raw data are smoothed with r-LOESS. Before replacing the holidays, the NN could see the holiday-patterns in the training data. But after replacing holidays in training data by non-holidays' data according to the Time-Window based technique, the NN always forecasts assuming that the next day is a non-holiday. This is the reason to get higher $MAPE_d^p$ on holidays.

May 25, 2013 (Saturday) is a special day. Even though it is not a holiday, it has a low electricity consumption compared to the other days. However, May 24, 2013 (Friday) is a holiday and that is the reason May 25, 2013 has a low electricity consumption. Therefore, after cleaned the training data with both Time-Window based and the r-LOESS methods, the NN forecasts electricity consumption assuming the next day is a regular non-holiday. This increases $MAPE_d^p$ and $MAPE_d^r$ from 2.8118 to 12.0632 and from 2.8118 to 11.8075, respectively.

When compared the results between the Time-Window based data cleaning technique and r-LOESS, the total average MAPE is lower with r-LOESS since its holidays' average MAPE is lower compared to that MAPE with the Time-Window based data cleaning technique. Except weekends' average MAPE, all the other categories under non-holidays' MAPE are better with the Time-Window based data cleaning technique.

TABLE III

MAPE table for all testing days of May 2013

| Date | Day | Holiday (Yes/No) | Bridging Holiday (Yes/No) | Before Cleaning the Data | Data cleaning with Time-Window based technique | Data cleaning with r-LOESS | Total load (MWs) | Peak load (MWs) |
|---|---|---|---|---|---|---|---|---|
| | | | | $MAPE_d$ | $MAPE_d^p$ | $MAPE_d^r$ | | |
| 5/1/2013 | Wed | Yes | No | 21.7383 | 26.0934 | 19.2670 | 296,571.70 | 7,322.55 |
| 5/2/2013 | Thu | No | No | 6.6505 | 5.4363 | 5.6952 | 363,797.25 | 9,537.35 |
| 5/3/2013 | Fri | No | No | 6.7731 | 7.3694 | 6.2955 | 387,590.85 | 9,701.85 |
| 5/4/2013 | Sat | No | No | 3.9711 | 6.1252 | 3.4097 | 359,577.75 | 8,437.05 |
| 5/5/2013 | Sun | Yes | No | 5.6705 | 6.1357 | 3.9256 | 290,664.45 | 7,345.95 |
| 5/6/2013 | Mon | Yes | No | 5.1747 | 8.3640 | 8.5274 | 323,167.15 | 7,926.00 |
| 5/7/2013 | Tue | No | No | 3.3561 | 3.1451 | 4.4690 | 367,490.50 | 9,507.30 |
| 5/8/2013 | Wed | No | No | 6.9022 | 4.9940 | 3.7939 | 380,694.75 | 9,550.95 |
| 5/9/2013 | Thu | No | No | 4.2302 | 4.5996 | 5.3194 | 381,441.25 | 9,494.70 |
| 5/10/2013 | Fri | No | No | 4.2388 | 5.0188 | 3.1330 | 381,153.45 | 9,595.35 |
| 5/11/2013 | Sat | No | No | 3.3115 | 1.6169 | 2.4811 | 355,820.90 | 8,392.10 |
| 5/12/2013 | Sun | No | No | 1.9381 | 2.1589 | 2.1669 | 309,370.50 | 7,369.85 |
| 5/13/2013 | Mon | Yes | No | 2.6903 | 2.7093 | 3.3696 | 359,441.65 | 8,876.80 |
| 5/14/2013 | Tue | No | No | 5.4659 | 4.8397 | 5.5129 | 385,717.65 | 9,647.85 |
| 5/15/2013 | Wed | No | No | 5.8322 | 4.3823 | 3.9455 | 393,716.25 | 9,863.95 |
| 5/16/2013 | Thu | No | No | 3.1152 | 4.4602 | 2.6372 | 395,520.63 | 10,013.35 |
| 5/17/2013 | Fri | No | No | 2.4634 | 3.7614 | 3.0470 | 393,245.39 | 9,869.10 |
| 5/18/2013 | Sat | No | No | 3.1881 | 4.5823 | 4.1989 | 353,273.55 | 8,190.75 |
| 5/19/2013 | Sun | No | No | 5.8933 | 5.0643 | 4.8445 | 294,841.75 | 7,054.20 |
| 5/20/2013 | Mon | No | No | 3.7583 | 3.0981 | 3.9825 | 353,309.25 | 9,305.40 |
| 5/21/2013 | Tue | No | No | 4.5056 | 2.5373 | 3.4726 | 369,124.30 | 9,380.60 |
| 5/22/2013 | Wed | No | No | 2.8888 | 2.4395 | 3.6217 | 364,559.50 | 9,332.60 |
| 5/23/2013 | Thu | No | No | 2.5068 | 2.4525 | 4.4883 | 362,643.35 | 9,360.00 |
| 5/24/2013 | Fri | Yes | No | 9.6774 | 15.9066 | 13.5930 | 313,282.05 | 7,262.55 |
| 5/25/2013 | Sat | No | No | 2.8118 | 12.0632 | 11.8075 | 312,171.45 | 7,423.50 |
| 5/26/2013 | Sun | No | No | 6.9823 | 4.1106 | 4.1985 | 287,719.05 | 7,178.20 |
| 5/27/2013 | Mon | No | No | 5.2878 | 2.0680 | 3.5238 | 351,069.15 | 9,278.10 |
| 5/28/2013 | Tue | No | No | 4.6061 | 2.7756 | 3.9788 | 373,677.00 | 9,456.45 |
| 5/29/2013 | Wed | No | No | 3.8512 | 1.3801 | 2.4095 | 376,834.90 | 9,508.60 |
| 5/30/2013 | Thu | No | No | 2.5948 | 3.3701 | 3.3582 | 368,693.38 | 9,224.75 |
| 5/31/2013 | Fri | No | No | 9.1448 | 8.5141 | 6.1866 | 347,458.50 | 8,963.65 |

TABLE IV
Average MAPEs different day categories in May 2013

| | Before | Data cleaning with Time-Window based technique | Data cleaning With r-LOESS |
|---|---|---|---|
| Total average | 5.2006 | 5.5346 | 5.1826 |
| Holidays' average | 8.9902 | 11.8418 | 9.7366 |
| Non-Holidays' average | 4.5383 | 4.0120 | 4.3069 |
| Mondays' average | 4.5231 | 2.5830 | 3.7532 |
| Weekdays' (Tue–Fri) average | 4.6810 | 4.1798 | 4.1980 |
| Weekends' average | 4.2141 | 3.9430 | 3.5500 |
| May 25th | 2.8118 | 12.0632 | 11.8075 |

## VIII. Conclusion

The main objective of this research is to propose an effective data preprocessing step for electricity demand forecasting. Thefocus is on reducing the forecasting errors on non-holidays as it is one of the most important phases in the short-term load forecasting. The NN outputs with raw data and data which arecleaned with the r-LOESS technique are set as the benchmark to compare the NN outputs with cleaned data where holidays, bridging holidays, and outliers were replaced with the Time-Window based technique. MAPEs are calculated for all three categories, separately.

The testing data (May, 2013) can be divided into two groups as holidays and non-holidays. All the results for non-holidays convene to reduce the average MAPE both with the Time-Window based technique and r-LOESS. As the Time-Window based preprocessing phase increases the performances of forecasting outcomes by 11.60% for non-holidays, we strongly recommend this preprocessing step for electricity demand forecasting.

For increasing the overall forecasting performances by decreasing the forecasting errors on holidays and bridging holidays, this research can be extended further with a new preprocessing stage. The concept of creating the filtering band should be changed and new outlier replacement techniques has to be used. At the same time, use of secondary experiments such as trial and error methods within the research has to be reduced in order to decrease the total research or experiment times. However, in this research, we had to conduct several secondary experiments to bring the best overall outcomes. But the use of statistical methods and optimizations would reduce the number of experiments. These points can be performed and improved in the future research for reducing the overall forecasting errors.

## Acknowledgment

This research is partially supported by the Logistics and Supply Chain Systems Engineering Research Unit (LogEn), Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU). Data used in this research is provided by Electricity Generating Authority, Thailand (EGAT). Therefore, we acknowledge their support for completing this research, successfully.

## Reference

[1] A. D. Papalexopoulos, and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Transactions on Power Systems,* vol. 5, no. 4, pp. 1535-1547, 1990.

[2] R. Sadownik, and E. P. Barbosa, "Shortterm forecasting of industrial electricity consumption in Brazil," *Journal of Forecasting,* vol. 18, no. 3, pp. 215-224, 1999.

[3] Y. Ohtsuka, and K. Kakamu, "SpaceTime Model versus VAR Model: Forecasting Electricity demand in Japan." *Journal of Forecasting,* vol. 32, no. 1, pp. 75-85, 2013.

[4] U. D. Caprio et al., "Short term load forecasting in electric power systems: a comparison of ARMA models and extended wiener filtering." *Journal of Forecasting,* vol. 2, no. 1, pp. 59-76, 1983.

[5] G. J. Janacek, and L. Swift, *Time series: forecasting, simulation, applications.* Ellis horwood, 1993.

[6] A. K. Singh et al., "An overview of electricity demand forecasting techniques. *Network and Complex Systems,"* vol. 3, no. 3, pp. 38-48, 2013.

[7] M. H. Amini, A. Kargarian, and O. Karabasoglu, "ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation." *Electric Power Systems Research,* vol. 140, pp. 378-390, 2016.

[8] S. Voronin, and J. Partanen, "Forecasting electricity price and demand using a hybrid approach based on wavelet transform," *ARIMA and neural networks. International Journal of Energy Research,* vol. 38, no. 5, pp. 626-637, 2014.

[9] A. K. Topalli, I. Erkmen, and I. Topalli, "Intelligent short-term load forecasting in Turkey." *International Journal of Electrical Power & Energy Systems,* vol. 28, no. 7, pp. 437-447, 2006.

[10] S. C. Bhattacharyya, and L. T. Thanh, "Shortterm electric load forecasting using an artificial neural network: case of Northern Vietnam." *International journal of energy research,* vol. 28, no. 5, pp. 463-472, 2004.

[11] F. Cavallaro, "Electric load analysis using an artificial neural network." *International journal of energy research,* vol. 29, no. 5, pp. 377-392, 2005.

[12] G. Nasr, E. Badr, and M. Younes, "Neural networks in forecasting electrical energy consumption: univariate and multivariate approaches." *International Journal of Energy Research,* vol. 26, no. 1, pp. 67-78, 2002.

[13] D. Singh, and S. Singh, "A self-selecting neural network for short-term load forecasting." *Electric Power Components and Systems,* vol. 29, no. 2, pp. 117-130, 2001.

[14] A. Demiroren, and G. Ceylan, "Middle anatolian region short-term load forecasting using artificial neural networks." *Electric Power Components and Systems,* vol. 34, no. 6, pp. 707-724, 2006.

[15] T. Senjyu, H. Sakihara, Y. Tamaki, and K. Uezato, "Next-day load curve forecasting using neural network based on similarity." *Electric Power Components and Systems,* vol. 29, no. 10, pp. 939-948, 2001.

[16] A. Dedinec et al., "Deep belief network based electricity load forecasting: An analysis of Macedonian case." *Energy,* vol. 115, pp. 1688-1700, 2016.

[17] S. Fan, L. Chen, and W.-J. Lee, "Machine learning based switching model for electricity load forecasting." *Energy Conversion and Management,* vol. 49, no. 6, pp. 1331-1344, 2008.

[18] J. C. Sousa, H. M. Jorge, and L. P. Neves, "Shortterm load forecasting based on support vector regression and load profiling." *International Journal of Energy Research,* vol. 38, no. 3, pp. 350-362, 2014.

[19] M. Mohandes, "Support vector machines for shortterm electrical load forecasting." *International Journal of Energy Research,* vol. 26, no. 4, pp. 335-345, 2002.

[20] D. Srinivasan, and M. Lee, "Survey of hybrid fuzzy neural approaches to electric load forecasting. in Systems, Man and Cybernetics," Intelligent Systems for the 21st Century, IEEE International Conference. 1995.

[21] S. Kazemi et al., "An evolutionary based adaptive neuro fuzzy inference system for intelligent shortterm load forecasting." *International Transactions in Operational Research,* vol. 21, no. 2, pp. 311-326, 2014.

[22] K. -L. Ho et al., "Short term load forecasting of Taiwan power system using a knowledge-based expert system." *IEEE Transactions on Power Systems,* vol. 5, no. 4, pp. 1214-1221, 1990.

[23] B. Kermanshahi, and H. Iwamiya, "Up to year 2020 load forecasting using neural nets." *International Journal of Electrical Power & Energy Systems,* vol. 24, no. 9, pp. 789-797, 2002.

[24] D. Bunn, and E. Farmer, "Review of short-term forecasting methods in the electric power industry." *Comparative models for electrical load forecasting,* pp. 13-30, 1985.

[25] H. K. Alfares, and M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods." *International Journal of Systems Science,* vol. 33, no. 1, pp. 23-34, 2002.

[26] I. Ben-Gal, *Outlier detection, in Data mining and knowledge discovery handbook.* Springer, 2005, pp. 131-146.

[27] Y. Yu et al., "Time series outlier detection based on sliding window prediction." *Mathematical Problems in Engineering,* 2014.

[28] S. Xu et al., "An improved methodology for outlier detection in dynamic datasets." *AIChE Journal,* vol. 61, no. 2, pp. 419-433, 2015.

[29] S. Hekimoglu, B. Erdogan, and R. Erenoglu, "A new outlier detection method considering outliers as model errors." *Experimental Techniques,* vol. 39, no. 1, pp. 57-68, 2015.

[30] C. Chen, and L. M. Liu, "Forecasting time series with outliers." *Journal of Forecasting,* vol. 12, no. 1, pp. 13-35, 1993.

[31] H. Liu, S. Shah, and W. Jiang, "On-line outlier detection and data cleaning." *Computers & chemical engineering,* vol. 28, no. 9, pp. 1635-1647, 2004.

[32] S. Trueck, R. Weron, and R. Wolff, *Outlier treatment and robust approaches for modeling electricity spot prices,* 2007.

[33] A. S. Hadi, "Identifying multiple outliers in multivariate data." *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 761-771, 1992.

[34] A. S., Hadi, "A modification of a method for the detection of outliers in multivariate samples." *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 393-396, 1994.

[35] J. H. Janssens, I. Flesch, and E.O. Postma. "Outlier detection with one-class classifiers from ML and KDD." in Machine Learning and Applications, ICMLA'09. International Conference. 2009.

[36] G. J. Williams et al., "A Comparative Study of RNN for Outlier Detection in Data Mining." in ICDM. 2002.

[37] S. Hawkins, et al., "Outlier detection using replicator neural networks." in International Conference on Data Warehousing and Knowledge Discovery. Springer, 2002.

[38] J. Chen et al., "Automated load curve data cleansing in power systems." *IEEE Transactions on Smart Grid,* vol. 1, no. 2, pp. 213-221, 2010.

[39] A. Tidemann, and H. Langseth, "Effects of data cleansing on load prediction algorithms." in Computational Intelligence Applications In Smart Grid (CIASG), IEEE Symposium. 2013.

[40] G. Tang et al., "From landscape to portrait: a new approach for outlier detection in load curve data." *IEEE Transactions on Smart Grid,* vol. 5, no. 4, pp. 1764-1773, 2014.

[41] R. M. do Nascimento et al., "Outliers' detection and filling algorithms for smart metering centers." in Transmission and Distribution Conference and Exposition (T&D), IEEE PES. 2012.

[42] G. Mateos, and G. B. Giannakis, "Load curve data cleansing and imputation via sparsity and low rank." *IEEE Transactions on Smart Grid,* vol. 4, no. 4, pp. 2347-2355, 2013.

[43] V. J. Hodge, and J. Austin, "A survey of outlier detection methodologies." *Artificial intelligence review,* vol. 22, no. 2, pp. 85-126, 2004.

[44] R. Rustum, and A. J. Adeloye, "Replacing outliers and missing values from activated sludge data using Kohonen self-organizing map." *Journal of Environmental Engineering,* vol. 133, no. 9, pp. 909-916, 2007.

[45] H. R. Maier, and G. C. Dandy, "The use of artificial neural networks for the prediction of water quality parameters." *Water Resour Res,* vol. 32, no. 4, pp. 1013-1022, 1996.

[46] A. Adeloye, and A. De Munari, "Artificial neural network based generalized storage–yield–reliability models using the Levenberg–Marquardt algorithm." *Journal of Hydrology,* vol. 326, no. 1, pp. 215-230, 2006.

[47] I. L. MacDonald, and W. Zucchini, *Hidden Markov and other models for discrete-valued time series*, vol. 110. CRC Press, 1997.

[48] A. C. Harvey, *Forecasting, structural time series models and the Kalman filter.* Cambridge university press, 1990.

[49] Z. Zhang et al., "A study on the method for cleaning and repairing the probe vehicle data." *IEEE Transactions on Intelligent Transportation Systems,* vol. 14, no. 1, pp. 419-427, 2013.

[50] M. Batmend, and D. Perdukova, "Linear regression based real-time filtering." *Advances in Electrical and Electronic Engineering,* vol. 11, no. 6, pp. 487, 2013.

[51] J. Connor, "A robust neural network filter for electricity demand prediction." *Journal of Forecasting,* vol. 15, no. 6, pp. 437-458, 1996.

[52] R. Weron, "Modeling and Forecasting Electricity Loads." *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach,* pp. 67-100, 2006.

[53] W. S. Cleveland, S. J. Devlin, and E. Grosse, "Regression by local fitting: methods, properties, and computational algorithms." *Journal of econometrics,* vol. 37, no. 1, pp. 87-114, 1988.
[54] W. S. Cleveland, and S. J. Devlin, "Locally weighted regression: an approach to regression analysis by local fitting." *Journal of the American statistical association,* vol. 83, no. 403, pp. 596-610, 1988.
[55] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots." *Journal of the American statistical association,* vol. 74, no. 368, pp. 829-836, 1979.

**Chawalit Jeenanunta** is an associate professor of School of Management Technology (MT), Sirindhorn International Institute of Technology, Thammasat University, Thailand. He received a B.S. degree in Mathematics and Computer Science, and M.Sc. in Management Science from University of Maryland and he received his Ph.D. in Industrial and Systems Engineering from Virginia Polytechnic Institute and State University. His Research interests are in area of applications of operations research, simulation, large-scaled optimization and supply chain management.

**K. Darshana Abeyrathna** is a Ph.D. research fellow at the University of Agder, Norway. He completed his BSc in Mechatronics Engineering at AIT university, Thailand in 2015. Then he directly joined to the Big Data research group at Sirindhorn International Institute of Technology, Thammasat University, Thailand for his MSc studies and graduated in 2017. His research interests are in Artificial Neural Networks, Data Mining, Forecasting, Optimization, and Operations Research.

**M. H. M. R. Syamali Dilhani** is a lecturer of Faculty of Engineering, University of Ruhuna, Sri Lanka. She received her BSc (Mathematics) (Hons) degree from the University of Ruhuna, Sri Lanka and M.Eng. in Logistic and Supply Chain Systems Engineering from Sirindhorn International Institute of Technology, Thammasat University, Thailand (2016). Her research interests are optimization techniques, artificial neural networks, and mathematical modeling.

**Su Wutyi Hnin** is a Ph.D. Student in Management Technology at Sirindhorn International Institute of Technology (SIIT), Thammasat University. She received her Bachelor of Engineering (B.E) degree in Electrical Power from Technological University, Mandalay, Myanmar in 2013. In 2017, she graduated Master of Science (Engineering and Technology). She received the Excellent Foreign Student (EFS) scholarship award for her Master Degree and Ph.D. from Sirindhorn International Institute of Technology, Thammasat University, Thailand. Her research interests are Energy forecasting, Optimization, Support Vector Regression.

**Pyae Pyae Phyo** graduated Bachelor of Technology (B.Tech) and Bachelor of Engineering (B.E) majoring in electrical power engineering from the Technological University Mandalay, Myanmar in 2013 and 2014 respectively. In 2015, she received the Excellent Foreign Student (EFS) scholarship award from Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU), Thailand.She graduated Master of Science (Engineering and Technology) in 2017. She is a Ph.D. student in Management Technology. Her current research interests include Deep Learning, Deep Neural Networks, Recurrent Neural Networks, Long Short-Term Memory, Regression,and Forecasting.