Applied Energy 183 (2016) 327–339

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Electric vehicle charging demand forecasting model based on big data technologies

Mariz B. Arias [a,b], Sungwoo Bae [a,*]

[a] Department of Electrical Engineering, Yeungnam University, Gyeongsan-si 38541, South Korea
[b] Department of Electrical Engineering, University of Santo Tomas, España, Manila 1015, Philippines

HIGHLIGHTS

- An EV charging demand forecasting model with big data technologies is proposed.
- The forecasting model uses historical real-world traffic data and weather data.
- A battery charging starting time is determined by real-world traffic patterns.
- The presented model considers charging demand for electric cars and buses.
- The proposed model considers both slow and fast charging classifications.

ARTICLE INFO

Article history:
Received 2 May 2016
Received in revised form 25 July 2016
Accepted 14 August 2016

Keywords:
Electric vehicle charging demand forecasting model
Big data
Real-world traffic data
Weather data
Cluster analysis

ABSTRACT

This paper presents a forecasting model to estimate electric vehicle charging demand based on big data technologies. Most previous studies have not considered real-world traffic distribution data and weather conditions in predicting the electric vehicle charging demand. In this paper, the historical traffic data and weather data of South Korea were used to formulate the forecasting model. The forecasting processes include a cluster analysis to classify traffic patterns, a relational analysis to identify influential factors, and a decision tree to establish classification criteria. The considered variables in this study were the charging starting time determined by the real-world traffic patterns and the initial state-of-charge of a battery. Example case studies for electric vehicle charging demand during weekdays and weekends in summer and winter were presented to show the different charging load profiles of electric vehicles in the residential and commercial sites. The presented forecasting model may allow power system engineers to anticipate electric vehicle charging demand based on historical traffic data and weather data. Therefore, the proposed electric vehicle charging demand model can be the foundation for the research on the impact of charging electric vehicles on the power system.

1. Introduction

Achieving sustainable transportation to address future energy requirements is now a vital mission of many countries. Electric vehicles (EVs) play a major role to increase energy security and to reduce emissions of greenhouse gases and other pollutants [1]. Because of insufficient domestic resources, South Korea relies on import to meet about 97% of its primary energy consumption [2]. The transportation sector was estimated to account for 17.9% of the country's energy consumption in 2015. To overcome the dependence on oil in this sector, the government has encouraged the development and usage of EVs and related infrastructure in the country.

However, if the penetration of these EVs increases in the future, electricity demand on the power system is also expected to increase because of charging their batteries from the utility grids. According to [3], large penetration of EVs may improve the sustainability of transportation but could also introduce various problems. Aside from the construction of related infrastructure, the main issue is the increase in electricity demand. Charging the battery of a single EV can increase household electricity consumption by 50% [3]. Therefore, several studies [4–12] have assessed the impact of charging EVs on the distribution systems. Various models have been developed to determine the impact of this additional EV charging demand to the distribution systems. In [8–12], a probabilistic approach was used to consider the factors in determining

* Corresponding author.
E-mail address: sbae@yu.ac.kr (S. Bae).

http://dx.doi.org/10.1016/j.apenergy.2016.08.080

the EV charging behaviors. In [11], a stochastic model of an EV load with the beginning battery charging time and the initial state-of-charge (SOC) of the battery was formulated to analyze its impact on the distribution system, and a comparative analysis on the four EV charging scenarios was also carried out. On the other hand, a recent study [12] concluded that EV charging demand variability is significant only in scenarios without control on charging EVs, but not in scenarios with control cases such as tariff controlled and smart charging scenarios, which reduce the EV charging demand variability. Although the research work presented in [12] considered EV model characteristic, mobility pattern, charging processes, social and economic variables, it did not consider the real-world data of weather and traffic volumes which may change the EV charging demand.

The prediction of EV charging demand is the foundation for the research on the impact of charging EVs on the power system. A number of methods have been proposed for the prediction of EV charging demand [13–25]. Some studies [13,14] have used a mathematical model, and other studies [5,6,15–19] have used different methods such as a Monte Carlo forecasting technique and Support Vector Machines for predicting the charging demand of EVs. In [18], a Monte Carlo simulation was used to obtain the spatial-temporal distribution of EVs considering the uncertainties of battery characteristics and transportation behaviors. A recent study [19] also presented different forecasting methods based on historical charging data; these datasets include charging records from customer profiles which are prone to privacy invasion issues, and station measurement records which contain a large volume of data to require long processing time. On the other hand, this paper presents an EV charging demand forecasting method which considers real-world traffic volume data with weather conditions that can resolve the privacy invasion issues and data processing speed concerns from the previous research works. Different variables were also considered in determining the electric vehicle charging demand. In [20], vehicle speed, acceleration, and roadway grade were used to model EV energy consumption. However, most studies considered driving patterns for predicting EV charging demand. The considered variables based on these driving patterns were starting charging time, charging periods, traveled distance, and battery SOC. In [21], the driving patterns were based on a national household travel survey, and studies in [22–24] used real-world driving patterns. A study [25] that statistically modeled the temporal availability of EVs charging by the utility grid considered travel survey data to generate travel patterns with a kernel density estimation method. Compared with the research works that used the travel survey data, this paper considers the real-world traffic volume with weather conditions in determining the travel patterns which may affect the EV charging demand forecasting.

This paper proposes an EV charging demand forecasting method based on big data including real-world traffic distribution data and weather data collected in South Korea in every hour. Big data sets are not only in the form of large files that require long processing time, but also in the collection of numerous small files [26,27]. In [27], a short-term load forecasting model was developed based on big data technologies to handle large quantities of data including smart meter and weather data. This study used the big data technologies proposed in [27] to introduce an EV charging demand forecasting model with the real-world traffic distribution data and weather data. Specifically, the proposed forecasting model is based on the historical traffic distribution data and weather data in South Korea which were collected, stored and managed by big data technologies. These data were processed by the following analyses: (i) A cluster analysis was used to classify historical traffic distribution data into clusters which were grouped by their similarities. (ii) The weather data were used to identify influential factors affecting the traffic patterns using a relational analysis. (iii) A decision tree was used to establish the relationship between the formed clusters and the influential factors to develop classification criteria. These processes were used to formulate the forecasting model to predict the EV charging demand using the month and day to be forecasted, and the number of EVs for input parameters of an EV charging demand forecasting program.

This paper contributes to the field of the EV charging demand forecasting by applying and discussing the following technical aspects:

(i) This study introduces the effects of both the real-world traffic volume data and the weather conditions that change in predicting the EV charging demand. The considered variables and characteristics of the presented EV charging demand forecasting model include the battery charging starting time determined by the traffic patterns from the real-world traffic volume data, its initial SOC, the type of battery, and its charging characteristics with different charging power classifications. Compared to previous studies [5,6,11,15], this paper also determines the charging starting time from the real data. However, this paper determines the charging starting time not only based on the formed traffic pattern from the real-world traffic volume but also based on the effect of the weather on the traffic volume. Compared with the previous works [5,6,11,15] that considered the stochastic nature of the charging starting time, this paper assumes that the charging starting time follows a Gaussian distribution. This Gaussian distribution is commonly used as a probability density model because many real-world phenomena are normally distributed when their samples are large enough. Therefore, the Gaussian distribution can effectively describe the stochastic nature of the charging starting time of electric vehicles.

(ii) In addition, this paper considers not only electric cars but also electric buses in forecasting the EV charging demand. This paper considers both slow and fast charging classifications although previous studies have only considered either a slow charging classification [5,11,15] or a fast charging classification [13,14]. The slow charging classification is considered for electric cars, and the fast charging classification is considered for electric buses. Furthermore, this paper considers that electric cars are charged either at home or at the workplace while electric buses are charged either at bus stations or their respective parking lots which are equipped with fast chargers. Therefore, the proposed EV charging demand model can anticipate the EV charging load profiles in the residential and commercial sites.

The remainder of the paper is organized as follows: Section 2 introduces the technical architecture and methodology of the proposed EV charging demand forecasting model based on big data technologies. An EV charging demand forecasting model with the charging demand analysis is presented in Section 3. Example case studies with simulation results are presented in Section 4 to discuss the effectiveness of the proposed EV charging demand forecasting model. Finally, Section 5 concludes the paper with summaries and findings.

2. Technical architecture and methodology of an EV charging demand forecasting model with big data technologies

This study developed an EV charging demand forecasting program of which technical design architecture is based on MATLAB built-in functions. The technical architecture of this program has four layers including data sources, data storage, data management, and data processing as shown in Fig. 1. As depicted in Fig. 1, the first layer covers the data sources that are stored by a local disk on computer. These data are collections of tabular text files which are composed of the historical traffic volume data and weather data of different roads in South Korea collected in every hour. The next layer is the data storage which is programed by the datastore function in MATLAB. This function is used to access the collections of data from the first layer. It can provide data in a chunk-wise manner for quick and efficient access to the target data. This function stores the traffic data and weather data of the specific road. The third layer manages the data stored in the previous layer. The MapReduce function in MATLAB which has the ability to perform calculations on large collections of data is used as the basis of this method. This function has the following three phases: map phase, intermediate phase, and reduce phase. The chunk of data enters the map phase which configures these data for processing. Then, the intermediate data go through the reduce phase which combines the intermediate results to generate a final result [28]. This study only used the map function that structures the stored traffic volume data and weather data into the desired format for data processing. The intermediate phase and the reduce phase were not included in this study because only a specific road was used in case studies for anticipating EV charging demand in Section 4 so that there were no similar data to be merged. This simplification for a specific road is not because of the forecasting capability of the proposed EV charging demand model but because of the limited space in this paper. However, these two phases are required to be included in forecasting the EV charging demand in more complex road cases which are out of the scope in this paper. The row of the data includes the date of datum, and the column of the data is the time of datum. The data consist of traffic volume, temperature, humidity, wind speed, and day types. The last layer processes these data to formulate an EV charging demand forecasting model. This last layer includes the following technologies: a cluster analysis, a relational analysis, and a decision tree. Similar to the methodology used in [27], the first step of the data processing is to apply a machine learning technique to identify typical traffic patterns of historical traffic volume data. The next step is to find influential factors in order to establish a decision tree to formulate the EV charging demand forecasting model.

2.1. Cluster analysis

The traffic volume consists of different traffic patterns resulted from various factors such as weather and day types. In this study, a hierarchical clustering technique was used to classify historical traffic data into several patterns. Specifically, this study used an agglomerative hierarchical method [29] which represents a bottom-up approach. This study initially gathered as many clusters as samples which were daily traffic volume for a particular route in South Korea collected in every hour for a year. The most similar traffic patterns were first merged as initial clusters among these

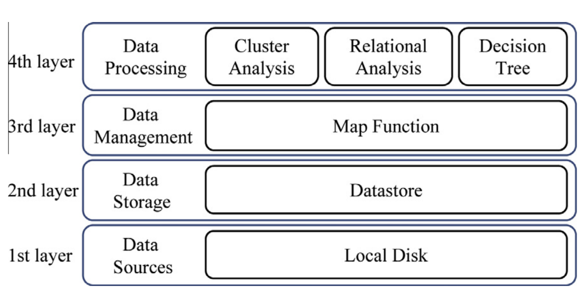

Fig. 1. Technical architecture of EV charging demand forecasting program with big data technologies.

samples. To establish their similarities, the pdist function in MATLAB was used in this study. This function computes the Euclidean distance between pairs of samples in the input matrix p, which is composed of row $m(1 \times n)$ vectors $x_1, x_2, \ldots, x_m$. The Euclidean distance between the ith and jth observations (i.e., $d(x(i,j))$ is defined in [29] as:

$$d(x(i,j)) = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2}. \tag{1}$$

The cophenetic correlation coefficient, which is the correlation between the cophenetic distance matrix and the original matrix, was determined to know how well a hierarchical clustering fitted the data. If the cophenetic correlation coefficient is closest to 1, it can be interpreted that the clustering technique is the most accurate. This paper used an average linkage algorithm which treats the distance between two clusters as an average distance between all pairs of samples where a member of a pair belongs to each cluster [30].

2.2. Relational analysis

Traffic volume is affected by many factors such as weather and day types. A grey relational analysis was used to discover the relationship of these factors to the traffic volume in this study. This grey relational analysis can determine important factors that have a significant influence on the traffic volume. The grey relational analysis generally consists of the following two main steps: The first step is to calculate the correlation of each point, and the next step is to calculate the correlation degree of each comparative series to the reference series based on an arithmetical mean. However, data pre-processing is normally required when the unit in one series is different from others. The data should be normalized into a comparable series before proceeding with the aforementioned two main steps. The original reference series was represented by $\boldsymbol{x}_0(\boldsymbol{k})$, and the original comparative series was denoted by $\boldsymbol{x}_i(\boldsymbol{k})$. In this study, the traffic volume was used as the reference series, and the factors affecting the traffic volume were used for the comparative series. After the data pre-processing, the next step is to calculate the grey relational coefficient ($\xi_{0i}(k)$) between two series by [31]

$$\xi_{0i}(k) = \frac{\Delta \min + \zeta \Delta \max}{\Delta_{0i} + \zeta \Delta \max}, \tag{2}$$

where $\Delta_{0i} = ||\boldsymbol{x}_0(\boldsymbol{k}) - \boldsymbol{x}_i(\boldsymbol{k})||$, $\Delta \min = \min_i \min_k ||\boldsymbol{x}_0(\boldsymbol{k}) - \boldsymbol{x}_i(\boldsymbol{k})||$, $\Delta \max = \max_i \max_k ||\boldsymbol{x}_0(\boldsymbol{k}) - \boldsymbol{x}_i(\boldsymbol{k})||$, and ζ is an identification coefficient which satisfies the following condition, $\zeta \in [0, 1]$. In this study, 0.5 was used for ζ because of its adequate distinguishing effect and stability [31].

After the grey relational coefficient ($\xi_{0i}(k)$) was calculated, the grey relational grade (γ_{0i}) can be determined by the average value of all the grey relational coefficients as follows [31]:

$$\gamma_{0i} = \frac{1}{n} \sum_{k=1}^{n} \xi_{0i}(k). \tag{3}$$

The grey relational grade (γ_{0i}) is the level of correlation between the reference series and the comparative series [31]. In this study, the grey relational grade was used to indicate the degree of the factor's influence on the traffic volume. The closer the grey relational grade is to 1, the more effect the factor has on the traffic volume. Although a factor of which grey relational grade was less than 0.6 was considered to have a negligible influence on the traffic volume, a factor of which grey relational grade was greater than 0.6 was considered as an influential factor on the traffic volume.

2.3. Decision tree

To establish the relationship between the formed clusters of traffic patterns and influential factors, this study used a decision tree that is a classification method to predict responses to data. The responses of the decision tree denote the formed clusters of cars and buses by the cluster analysis, and the data of the decision tree indicate the influential factors identified by the relational analysis. This decision tree that consists of nodes (i.e., decisions) and leaf nodes (i.e., responses) can be generated by repeatedly partitioning data set until the terminal node (i.e., stopping criterion) that only results in one response is reached. To predict such response, decisions in the tree should be made from the top node down to a leaf node. In this study, a fitctree function in MATLAB was used to construct the decision tree [28]. Based on the decision trees in the Appendices, the data of the influential factors determine the specific car and bus clusters in which the forecasting day belongs. Once the clusters are obtained, the EV charging demand can be estimated by the analysis presented in the proceeding section.

3. EV charging demand forecasting model

3.1. Considered charging factors for an EV charging demand analysis

EV charging demand is usually influenced by various factors such as EV battery charging characteristics, charging starting time, and charging power classification. These three factors which have effects on the charging load profiles of EVs were considered to analyze the EV charging demand in this study.

3.1.1. EV battery charging characteristics

In order to determine the charging demand of EVs, this study considered EV battery charging characteristics including vehicle types, battery technology, battery capacity, and initial battery SOC before it is charged. In this paper, vehicles were classified by two types including a car and a bus. A car denotes a private light-duty vehicle, and a bus represents a commercial heavy-duty passenger vehicle. The battery capacity may vary by the type and size of a vehicle. This study assumed that a car used a lithium-ion battery with a range of 148 km and a capacity of 27 kW h modeled by a Soul EV manufactured by Kia Motors. A bus was assumed to use a lithium-ion polymer battery which has a capacity of 85.8 kW h and a range of 69.8 km modeled by a Hankuk Fiber E-Primus bus. These two EV types were selected because they have been developed and deployed in Korea [32]. According to Korea Transportation Safety Authority [33], the average daily traveled mileages of a car and a bus are 32.6 km and 206.9 km respectively. Based on these average daily mileages, this study assumed that a car can be charged once a day although a bus should be charged two or three times a day during the off-peak hours depending on traffic patterns and bus clusters. This assumption was based on the fact that the average daily traveled mileage of an electric bus exceeds its driving range with a single charge per day.

An initial battery SOC before its charging has also an influence on EV charging demand. This study assumed that the initial battery SOC of an electric car before its charging was a random variable with a Gaussian distribution of which the probability density function (*pdf*) is defined as [34]:

$$f(s) = \frac{1}{\sigma_s\sqrt{2\pi}} e^{\frac{-(s-\mu_s)^2}{2\sigma_s^2}}, \tag{4}$$

where s is an initial battery SOC, μ_s is an average initial SOC value, and σ_s is its standard deviation. The initial SOC of an electric car battery was assumed to be between 0.2 and 0.8 because of the protecting algorithm of a battery management system to optimize the battery's lifetime. Therefore, the initial SOC of a car before its charging was generated by the random sampling of a Gaussian *pdf* with $\mu_s = 0.5$ and $\sigma_s = 0.3$ [34]. This study also assumed that an electric bus is immediately charged at bus stations or its parking lots which are equipped with fast chargers whenever the SOC of its battery reaches 0.20 to optimize battery's life expectancy [34].

3.1.2. Charging starting time

This study determined an EV charging starting time based on the daily transportation behavior from traffic patterns classified by the cluster analysis. In this study, a battery charging for an electric car was assumed to begin when an electric car arrived at the workplace in the daytime or at home in the evening. This study assumed that an electric bus can be charged at bus stations or its respective parking lots which are equipped with fast chargers whenever the SOC of its battery reaches 0.2 during the off-peak traffic hours. The charging starting time for these vehicles depends on the traffic pattern observed on each cluster classified by the cluster analysis. Moreover, for the charging starting time this study used a random variable of which distribution is a Gaussian *pdf* for each EV cluster, which is specifically discussed in Section 3.6.

3.1.3. Charging power classification

The EV charging demand is also related to its charging power classification which determines the charging speed and charging power of an EV. There are three charging classifications according to the Society Automotive Engineers (SAE) Standard J1772 [35]. However, this study only used slow and fast charging power presented in Table 1 among the three charging classifications. In other words, an electric car was assumed to be charged by the level 1 charging power (i.e., slow charging) because an electric car can be charged at home or workplace for a long period of time. On the other hand, the level 3 charging power (i.e., fast charging) was considered for an electric bus because of its long daily mileage and limited charging time. This study also used a battery charging profile which was based on the simplified piecewise-linear charging profile model used in [36]. Fig. 2 shows the battery charging power and the related battery SOC profiles used for an electric car (i.e., Fig. 2(a)) and an electric bus (i.e., Fig. 2(b)).

3.2. Description of traffic data set

In order to formulate an EV charging demand forecasting model, this paper used the historical real traffic data of the entire South Korea collected by the Traffic Monitoring System (TMS) of the Ministry of Land, Infrastructure and Transport (MOLIT) [37]. These traffic data include daily traffic volumes and vehicle mileages by vehicle types collected in the highway, national route, and local roads of South Korea in every hour. To formulate the proposed EV charging demand forecasting model, this study used the training data that consist of the traffic volume and weather data observed on a national route particularly from Goyang to Paju captured in every hour from January 1, 2014 to December 31, 2014. In addition, although the current EV traffic volume is extremely less than the conventional vehicle traffic volume, this study assumed that EV traffic patterns will be identical with the real traffic data of conventional vehicles because EV penetration grows to signifi-

Table 1
Recommended charging classification.

Level	Classification	Charging power (kW)
Level 1	Slow (Home)	3.7
Level 1	Slow (Workplace)	3.7
Level 3	Fast	115

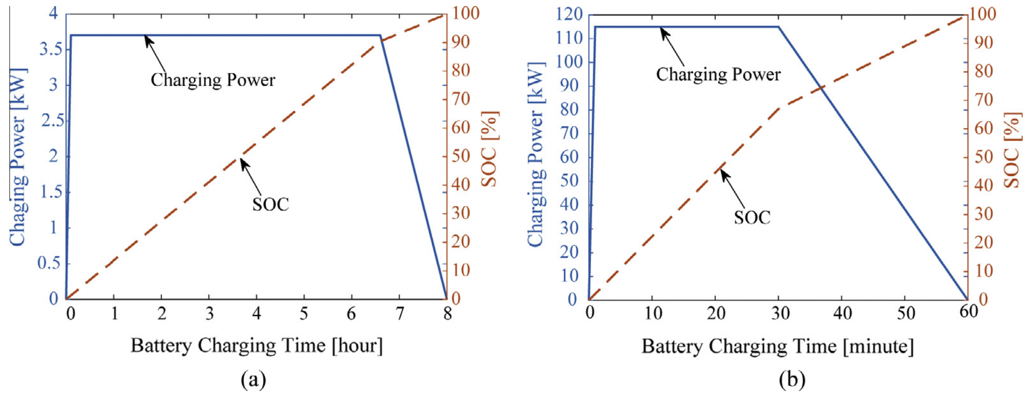

Fig. 2. Battery charging profiles of a Li-ion battery: (a) (Left) Electric car. (b) (Right) Electric bus.

cant penetration levels in the future. Figs. 3 and 4 respectively show the traffic volume distribution of conventional cars and buses in the chosen route from 1/1/2014 to 12/31/2014 obtained by the first three layers of the technical architecture of the EV charging demand forecasting program described in Section 2.

3.3. Classification of traffic patterns by a cluster analysis

As shown in Figs. 5(a)–(d) and 6(a)–(c), four clusters of cars and three clusters of buses were identified from the vehicle distribution data set by the aforementioned hierarchical clustering algorithm in Section 2.1. Table 2 summarizes the number of days associated with each cluster for cars and buses. A car cluster 1 (i.e., C1) is composed of event days such as festivals and international conferences. A car cluster 2 (i.e., C2) contains weekdays and most Saturdays. A car cluster 3 (i.e., C3) contains holidays. A car cluster 4 (i.e., C4) is mostly composed of Sundays. For bus clusters, a bus cluster 1 (i.e., B1) contains most weekends. A bus cluster 2 (i.e., B2) contains most weekdays. A bus cluster 3 (i.e., B3) consists of event days such as international conferences. This cluster analysis considered the events and international conferences that happened in the area nearby the chosen route during the studied period.

3.4. Identification of influential factors using a relational analysis

The influential factors affecting the traffic volume of cars and buses were determined by the grey relational analysis in Section 2.2. Table 3 shows the grey relational grades of factors affecting the traffic volume of cars and buses. Based on the results in Table 3, the maximum temperature, average temperature, average humidity, and day type are more influential on the traffic volume of cars than average wind speed. For buses, the maximum temperature, average temperature, average humidity, and average wind speed have more influential effect on the traffic volume of buses than day type. Therefore, the maximum temperature, average temperature, average humidity, and day type were used with parameters for the decision tree analysis of cars. On the other hand, the maximum temperature, average temperature, average humidity, and average wind speed were used with parameters for the decision tree analysis of buses.

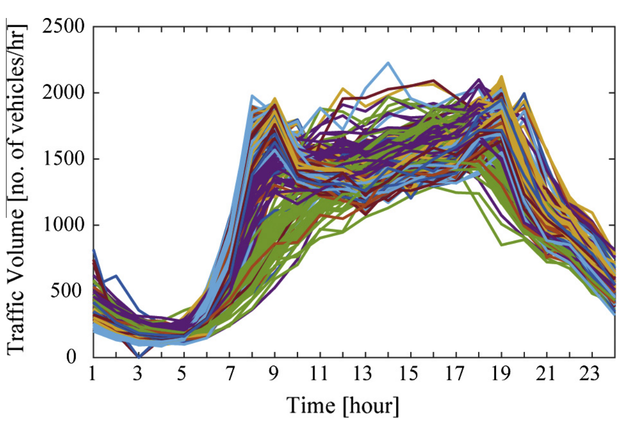

Fig. 3. Car distribution from Goyang to Paju.

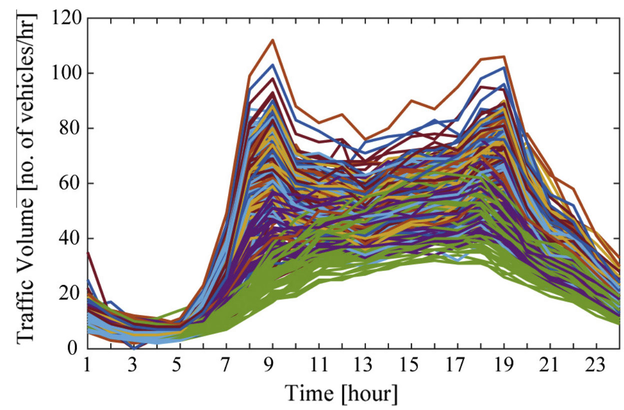

Fig. 4. Bus distribution from Goyang to Paju.

3.5. Establishment of a decision tree

The influential factors that affected the traffic volume of cars and buses and that were identified by the relational analysis were used to establish a decision tree described in Section 2.3. Using a fitctree function in MATLAB, the decision trees for cars and buses were established as shown in Figs. 9 and 10 in the Appendices. Specific car clusters (i.e., C1 $\sim$ C4) and bus clusters (i.e., B1 $\sim$ B3) can be determined with these decision trees if forecasted temperature, forecasted humidity, forecasted wind speed, and day type are given.

3.6. Determination of a charging starting time

Once the car and bus clusters were obtained from the decision trees with the forecasted weather data and day types, a charging starting time should be determined. To estimate the charging starting time of an EV battery, this study used the traffic patterns of the

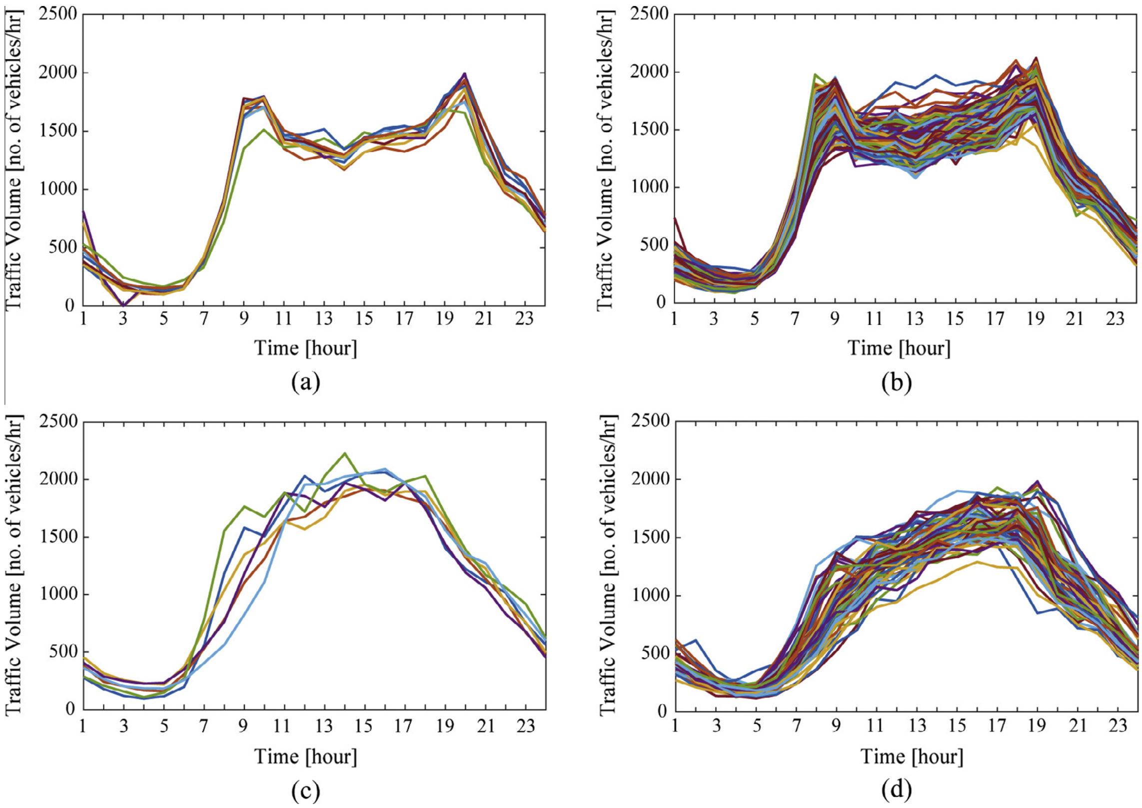

Fig. 5. Clustering results of cars from 1/1/2014 to 12/31/2014. (a) Car cluster 1. (b) Car cluster 2. (c) Car cluster 3. (d) Car cluster 4.

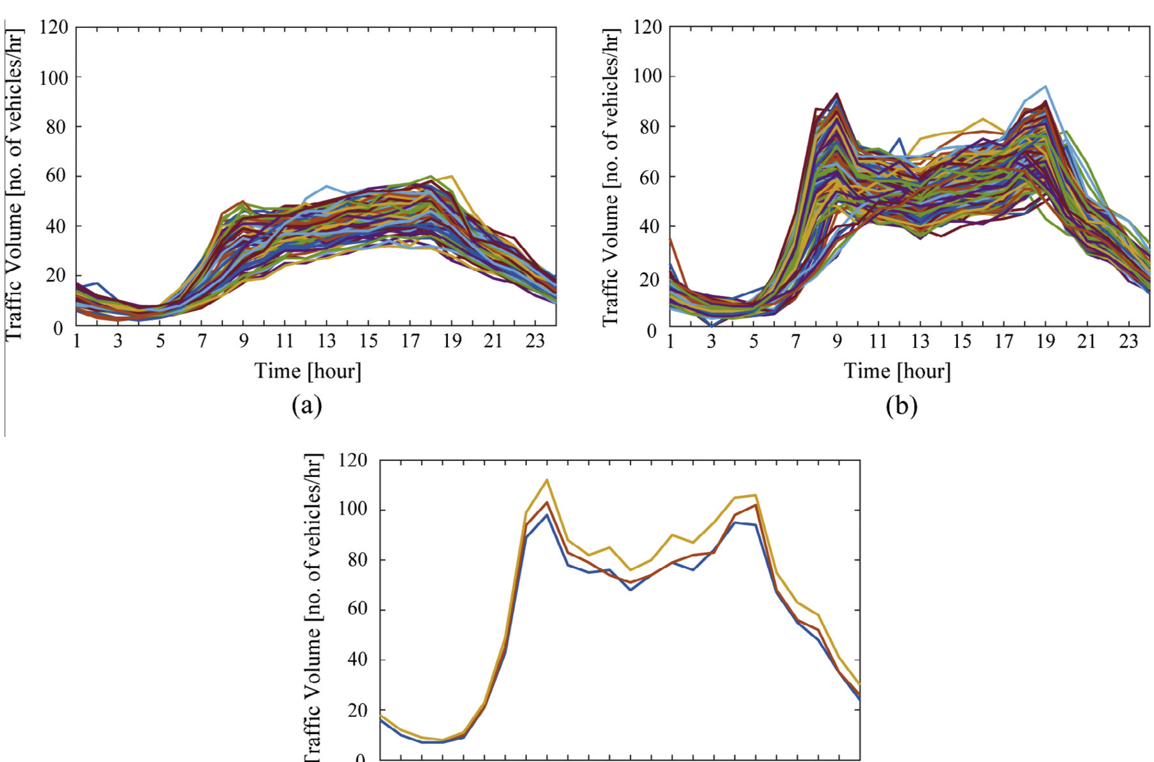

Fig. 6. Clustering results of buses from 1/1/2014 to 12/31/2014. (a) Bus cluster 1. (b) Bus cluster 2. (c) Bus cluster 3.

Table 2
Number of days associated with car and bus clusters.

Day type	C1	C2	C3	C4	B1	B2	B3
Monday	2	48	1	1	4	48	0
Tuesday	1	49	1	1	1	50	1
Wednesday	2	48	0	3	3	49	1
Thursday	2	47	0	3	2	49	1
Friday	3	47	1	1	2	50	0
Saturday	0	41	0	11	39	13	0
Sunday	0	0	3	49	47	5	0

Table 3
Grey relational grades of factors affecting the traffic pattern of cars and buses.

	MT	AT	AH	WS	DT
Cars	0.7585	0.7505	0.7512	0.5738	0.6264
Buses	0.6429	0.6525	0.7041	0.6819	0.5541

MT = maximum temperature, AT = average temperature, AH = average humidity, WS = average wind speed, DT = day type.

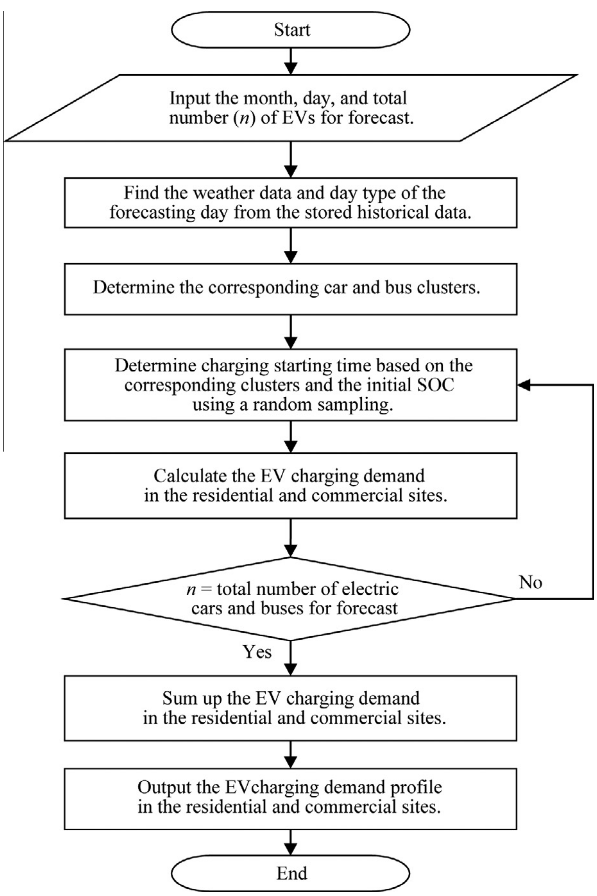

Fig. 7. Flowchart of the proposed EV charging demand forecasting model.

formed clusters acquired from the cluster analysis. Because this charging starting time of an EV battery is also a random variable discussed in Section 3.1.2, the Gaussian *pdf* of the charging starting time for each cluster was developed based on the traffic patterns observed in each cluster.

For car clusters, an electric car was assumed to be charged once a day because the average daily mileage of an electric car does not exceed its driving range with a single charge. In this study, an electric car can only be charged at the workplace in the daytime or at home in the night. To develop the Gaussian *pdf* of the charging starting time, the mean charging starting time and its standard deviation were determined by the traffic patterns formed in each cluster. There are two possible charging periods in these car clusters. The first charging period may occur in the daytime at the workplace. Because the morning highest peak traffic volume occurred in the morning rush hour and the minimum traffic volume in the daytime happened at the end of morning rush hour, the midpoint time of these two traffic volumes can be considered as the mean charging starting time in the daytime. The second charging period can happen at home in the evening. Because the evening highest peak traffic occurred in the evening rush hour, the time in which the traffic volume decreased by the half of the evening highest peak traffic volume was considered as the mean charging starting time in the evening. The typical charging starting time interval was used to determine the standard deviation that is defined by the half of its starting time interval [38]. Specifically, this study assumed that the typical charging starting time interval for electric cars is 2 h in the daytime because a typical electric car is charged once its user arrives at the workplace. On the other hand, the typical charging starting time interval for electric cars in the evening was assumed to be 4 h because an electric car user may not go to home immediately after he or she left from the office depending on his or her personal schedule.

3.6.1. Car cluster 1 (see Fig. 5(a))

This cluster is composed of event days such as festivals and international conferences. It was assumed that all cars in this cluster should be charged at home in the evening because high traffic volume was observed in the daytime. The Gaussian *pdf* of a charging starting time in the car cluster 1 (i.e., $f_{C1}(t)$) is given as:

$$f_{C1}(t) = \frac{1}{\sigma_{C1}\sqrt{2\pi}} e^{\frac{-(t-\mu_{C1})^2}{2\sigma_{C1}^2}}, \tag{5}$$

where t is a charging starting time, $\mu_{C1} = 23$ is the time in which the half of the evening highest peak traffic volume was observed, and $\sigma_{C1} = 2$ was used because the evening standard deviation for electric cars is the half of the evening charging starting time interval (i.e., 4 h) as aforementioned.

3.6.2. Car cluster 2 (see Fig. 5(b))

This cluster contains weekdays and most Saturdays. As shown in Fig. 5(b), two peak traffic volumes (i.e., the morning peak (between 8:00 and 9:00) and the evening peak (19:00)) were observed in this cluster. Because this cluster mainly consists of weekdays, it can be interpreted that there are two charging periods

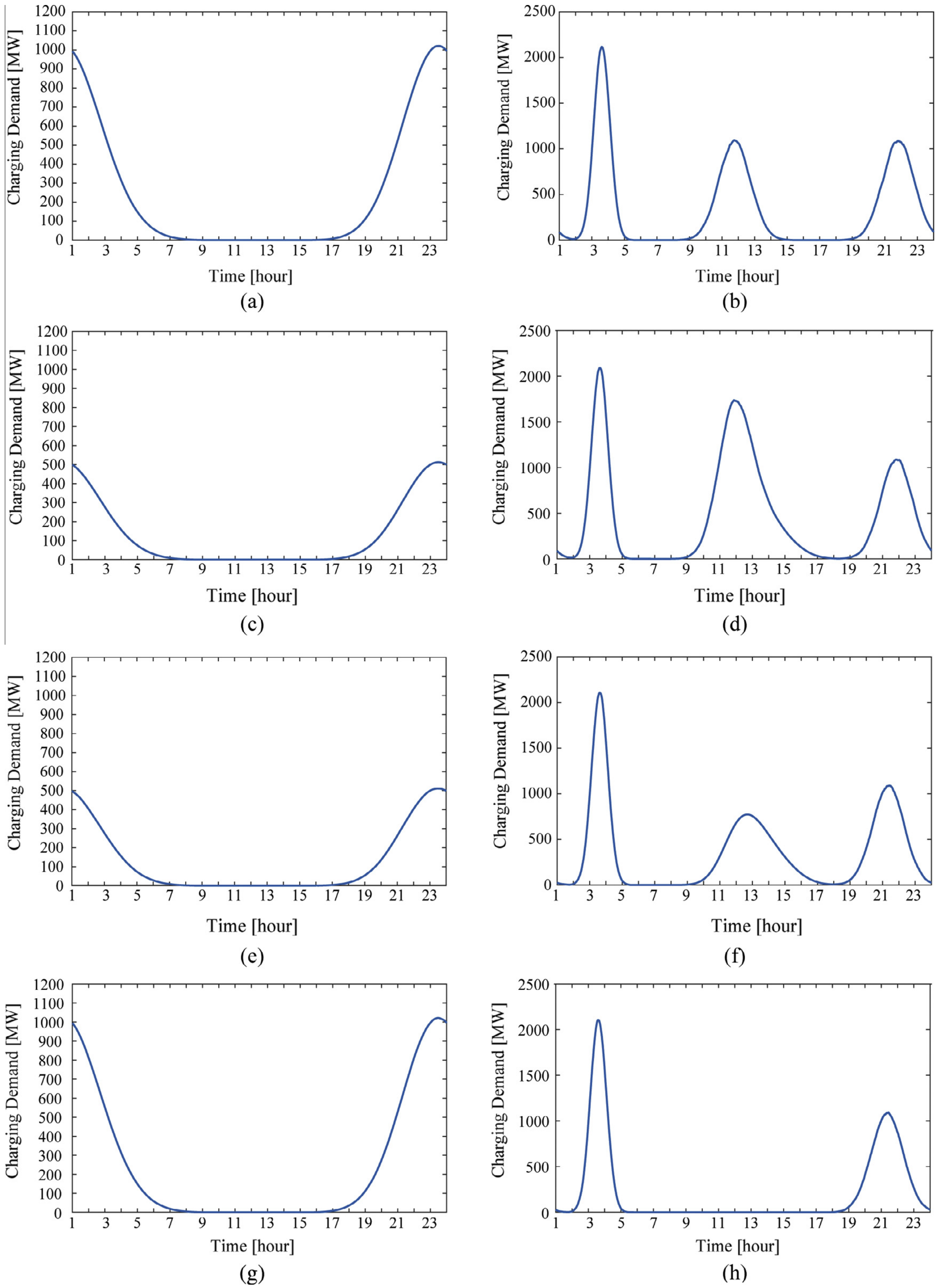

Fig. 8. Forecasted EV charging demand in the residential and commercial sites with the proposed EV charging demand model: (a) Winter/weekend/residential. (b) Winter/weekend/commercial. (c) Winter/weekday/residential. (d) Winter/weekday/commercial. (e) Summer/weekday/residential. (f) Summer/weekday/commercial. (g) Summer/weekend (also holiday)/residential. (h) Summer/weekend (also holiday)/commercial.

in this cluster. Similar to [11], this study assumed that 50% of electric cars are charged at the workplace in the daytime and that the other electric cars are charged at home in the evening. Based on this assumption, a charging starting time in the car cluster 2 is given by a Gaussian *pdf* (i.e., $f_{C2}(t)$) as follows:

$$f_{C2}(t) = \frac{0.5}{\sigma_{C2a}\sqrt{2\pi}} e^{\frac{-(t-\mu_{C2a})^2}{2\sigma^2_{C2a}}} + \frac{0.5}{\sigma_{C2b}\sqrt{2\pi}} e^{\frac{-(t-\mu_{C2b})^2}{2\sigma^2_{C2b}}}, \tag{6}$$

where $\mu_{C2a} = 11$ is the time in which the midpoint of the highest and minimum traffic volumes in the daytime was observed,

μ_{C2b} = 22 is the time in which the half of the evening highest peak traffic volume was observed, and σ_{C2a} = 1 and σ_{C2b} = 2 were used because the standard deviations are the half of the charging starting time intervals (i.e., 2 h in the daytime and 4 h in the evening) similar to the car cluster 1 case.

3.6.3. Car cluster 3 (see Fig. 5(c))

This cluster is composed of holidays. Because high traffic volume was observed in the daytime as shown in Fig. 5(c), all cars were assumed to be charged at home in the evening. The Gaussian *pdf* of a charging starting time in the car cluster 3 (i.e., $f_{C3}(t)$) can be described as follows:

$$f_{C3}(t) = \frac{1}{\sigma_{C3}\sqrt{2\pi}} e^{\frac{-(t-\mu_{C3})^2}{2\sigma_{C3}^2}}, \tag{7}$$

where μ_{C3} = 22 is the time in which traffic volume decreased by the half of the evening highest peak traffic volume, and σ_{C3} = 2 was used because of the same analysis with the car cluster 1.

3.6.4. Car cluster 4 (see Fig. 5(d))

This cluster mostly contains weekends. Because high traffic volume was observed in the daytime, cars were assumed to be charged at home in the evening. The Gaussian *pdf* of a charging starting time in the car cluster 4 (i.e., $f_{C4}(t)$) is described as:

$$f_{C4}(t) = \frac{1}{\sigma_{C4}\sqrt{2\pi}} e^{\frac{-(t-\mu_{C4})^2}{2\sigma_{C4}^2}}, \tag{8}$$

where μ_{C4} = 22 is the time in which the half of the evening highest traffic volume was observed, and σ_{C4} = 2 was used according to the same analysis with the car cluster 1.

3.6.5. Bus cluster 1 (see Fig. 6(a))

For bus clusters, all electric buses were assumed to be charged at their bus stations or respective parking lots as aforementioned. Because single charge per day cannot meet the average daily traveled mileages of an electric bus, this electric bus will be charged two or three times a day depending on the traffic patterns of each cluster. Hence, this study assumed that there are three charging periods in these clusters: in the daytime, in the evening, and in the non-operational hours (1:00–4:00). The mean charging starting time in these charging periods was determined by the midpoint between the times occurred in the highest peak and the minimum traffic volumes of each charging period. Similar to car clusters, the typical charging starting time interval for electric buses was used to determine the standard deviation of the charging starting time that is equal to the half of its interval [38]. Specifically, this study assumed that the charging starting time interval for bus clusters in the daytime and in the evening is 2 h because electric buses with various routes may be charged at different bus stations whenever their SOC reaches 0.2. In addition, the charging starting time interval for bus clusters is 1 h in the non-operational hours because electric buses are immediately charged at their parking lots at the end of the operation.

The bus cluster 1 is mainly composed of weekends. It was assumed that this bus cluster 1 has only two charging periods including evening and non-operational hours because there was high traffic volume in the daytime as show in Fig. 6(a). The Gaussian *pdf* of a charging starting time (i.e., t) in this bus cluster 1 (i.e., $f_{B1}(t)$) is described as follows:

$$f_{B1}(t) = \frac{1}{\sigma_{B1a}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B1a})^2}{2\sigma_{B1a}^2}} + \frac{1}{\sigma_{B1b}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B1b})^2}{2\sigma_{B1b}^2}}, \tag{9}$$

where μ_{B1a} = 2.5 was used for the non-operational hours because the midpoint of the highest and minimum traffic volumes in the non-operational hours was observed at this time, μ_{B1b} = 21 was determined because the midpoint of the evening highest and minimum traffic volumes was observed at this time, σ_{B1a} = 0.5 and σ_{B1b} = 1 were used because the standard deviations for bus clusters are equivalent to the half of the charging starting time intervals (i.e., 1 h in the non-operational hours and 2 h in the evening) as aforementioned. Contrary to Eq. (6), the numerators of each term in the right side of Eq. (9) should be one because an electric bus should be charged two or three times a day depending on the traffic patterns of each cluster as aforementioned.

3.6.6. Bus cluster 2 (see Fig. 6(b))

This cluster mostly contains weekdays. It was assumed that there were three charging periods in this cluster because two peak traffic volumes were observed as shown in Fig. 6(b). Therefore, the Gaussian *pdf* of a charging starting time in this bus cluster 2 (i.e., $f_{B2}(t)$) is given by:

$$f_{B2}(t) = \frac{1}{\sigma_{B2a}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B2a})^2}{2\sigma_{B2a}^2}} + \frac{1}{\sigma_{B2b}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B2b})^2}{2\sigma_{B2b}^2}} + \frac{1}{\sigma_{B2c}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B2c})^2}{2\sigma_{B2c}^2}}, \tag{10}$$

where μ_{B2a} = 2.5 was used for the non-operational hours because the midpoint of the non-operational highest and minimum traffic volumes was observed at this time, μ_{B2b} = 11 was used in the daytime charging starting time because the midpoint of the morning highest and minimum traffic volumes was observed at this time, and μ_{B2c} = 21.5 was used in the evening because the midpoint of the evening highest and minimum traffic volumes was observed at this time. In addition, σ_{B2a} = 0.5 and $\sigma_{B2b} = \sigma_{B2c}$ = 1 were used because the standard deviations for bus clusters are equal to the half of their respective charging starting time intervals (i.e., 1 h in the non-operational hours, and 2 h in the daytime and in the evening) as aforementioned. The numerators of each term in the right side of Eq. (10) should also be one as stated in the bus cluster 1.

3.6.7. Bus cluster 3 (see Fig. 6(c))

This is a special cluster which is composed of special event days such as international conferences. Because two peak traffic volumes were observed as shown in Fig. 6(c), it was assumed that there were three charging periods in this cluster. Therefore, the Gaussian *pdf* of a starting charging time in the bus cluster 3 (i.e., $f_{B3}(t)$) is described as:

$$f_{B3}(t) = \frac{1}{\sigma_{B3a}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B3a})^2}{2\sigma_{B3a}^2}} + \frac{1}{\sigma_{B3b}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B3b})^2}{2\sigma_{B3b}^2}} + \frac{1}{\sigma_{B3c}\sqrt{2\pi}} e^{\frac{-(t-\mu_{B3c})^2}{2\sigma_{B3c}^2}} \tag{11}$$

where μ_{B3a} = 2.5, μ_{B3b} = 11, μ_{B3c} = 21.5, σ_{B3a} = 0.5, and $\sigma_{B3b} = \sigma_{B3c}$ = 1 were used according to the same analysis with the bus cluster 2.

3.7. Development of an EV charging demand forecasting model

This section describes the overall development procedures of the EV charging demand forecasting model in this study. Once each car or bus cluster was identified from the decision trees as discussed in Sections 3.3–3.5, the charging starting time can be obtained by the random sampling of *pdfs* for each cluster as described in Section 3.6. The initial SOC before charging cars was also obtained by the random sampling as explained in Section 3.1.1. However, the initial SOC before charging buses was fixed to 0.20 as discussed in Section 3.1.1. In order to calculate the EV charging demand with computer software, this study used discretized charging power demand (P_d) and its related discretized SOC (SOC_d) from the charging profile as shown in Fig. 2. Once the initial SOC was obtained by the random sampling, the corresponding charging demand profile with the initial SOC can be plotted to a 24-h coordinate to determine single EV charging demand along a day. Although the single EV charging demand profile does not start from zero but from the determined initial SOC, it has the same shape of the battery charging profile as shown in Fig. 2. To obtain multiple EVs' charging demand, the aforementioned procedures were repeated until the total EV charging data were attained. The total charging demand of n number of EVs at any time t can be determined by:

$$P_t(t) = \sum_{i=1}^{n} P_d(t) \tag{12}$$

where $P_t(t)$ is the total EV charging demand at time t, n is the total number of EVs, and $P_d(t)$ is the charging demand power of an EV at time t. The charging demand profile of multiple EVs follows the Central Limit Theorem that a distribution approaches a normal distribution as the number of samples increases regardless of an actual distribution. As a result, the charging demand profile of multiple EVs shows a normal distribution because of the random variables including the charging starting time and the initial battery SOC.

Fig. 7 shows the flowchart of the proposed EV charging demand forecasting program. The inputs of the forecasting model include the month and day to be forecasted, and the total number of EVs. Once these inputs are passed to the program, the weather data and day types of the forecasting day can be found from the stored historical data by the computer program. Based on these weather data and day types with the decision trees, the corresponding car and bus clusters can be determined. Then, the charging starting time can be obtained according to the classified traffic patterns with the random sampling generated by the Gaussian *pdfs* of each cluster. The initial SOC of electric cars can also be determined by the random sampling as aforementioned in Section 3.1.1. The EV charging demand in the residential sites is calculated based on the charging demand of electric cars which are charged at home in the evening. On the other hand, the EV charging demand in the commercial sites is the sum of the charging demand of electric cars to be charged at the workplace in the daytime and the charging demand of electric buses. The outputs of this forecasting model include the charging load profiles in the residential and commercial sites.

4. Case studies and discussions

This section provides case studies with simulation results of the EV charging demand forecasting model presented in Section 3. The corresponding weather data from 1/1/2015 to 12/31/2015 were used for testing data. As listed in Table 4, four sample forecasted days for case studies were chosen to anticipate the EV charging demand of the weekdays and weekends in winter and summer. Once the forecasting temperature, forecasting humidity, wind speed and day types were identified, the car and bus clusters of the traffic volume was determined based on the decision trees in Figs. 9 and 10 in the Appendices. For example, the weather on 1/4/2015, Sunday, was the maximum temperature of 3 °C, an average temperature of −2 °C, an average humidity of 51%, and a wind speed of 9 km/h. According to the decision trees in Figs. 9 and 10 provided in the Appendices, car and bus clusters in this date belong to the car cluster 4 (i.e., C4) and the bus cluster 2 (i.e., B2) respectively. Based on the same analysis with the decision trees in Figs. 9 and 10, the corresponding clusters of cars and buses in those four forecasted day examples are listed in Table 4.

Table 4
Classification results.

Date	Day type	Car cluster	Bus cluster
1/4/2015	Weekend	C4	B2
1/22/2015	Weekday	C2	B2
8/6/2015	Weekday	C2	B1
8/15/2015	Weekend (also holiday)	C3	B1

In the case study examples, 1 million EVs were assumed in the simulation because they are the EV deployment goal in South Korea by 2020 [32]. Among the 20.12 million registered vehicles in South Korea in 2014, cars account for 78.3% of the total registered vehicles while the proportion of buses in the total registered vehicles are only 4.7% [39]. These statistical proportions of cars and buses in the conventional vehicles in 2014 were used for anticipating the charging demand of electric cars and buses in this simulation study. Based on these simulation conditions with the EV charging demand forecasting model discussed in Section 3, the EV charging demand was calculated in these four example days in winter and summer. In addition, this study assumed the following conditions: (1) electric cars will be charged once a day either at the workplace in the daytime or at home in the night with slow charging speed. On the other hand, electric buses will be charged two or three times a day either at their bus stations or at their respective parking lots with fast charging speed; (2) the charging starting time of electric cars and electric buses follows a Gaussian distribution based on the formed traffic patterns; and (3) the initial SOC of electric cars before its charging also follows a Gaussian distribution while the initial SOC of electric buses before its charging is 0.20.

Fig. 8 shows the charging demand load profiles of these forecasted days in winter and summer. As shown in Fig. 8, the charging demand was divided into two areas including residential and commercial sites. The charging demand for electric cars charged at home was considered as EV charging loads in the residential sites. On the other hand, the charging demand for electric cars charged at the workplace and for electric buses charged at their respective parking lots or bus stations accounted for EV charging loads in the commercial sites.

Fig. 8(a) and (b) shows the forecasted EV charging demand on a weekend in winter in the residential and commercial sites respectively. Fig. 8(a) shows high EV charging demand during the night because all cars were charged at home due to high traffic volume of electric cars on the road in the daytime on a weekend as depicted in Fig. 5(d). Fig. 8(b) shows only the charging demand of electric buses because no electric cars were charged at the workplace on a weekend. As depicted in Fig. 8(b), three charging load profiles were observed, and high charging demand was observed during the non-operational hours because there was only 1 h charging starting time interval during this period.

Fig. 8(c) and (d) shows the EV charging demand on a weekday in winter. Compared with Fig. 8(a), the charging load profile in the residential sites shown in Fig. 8(c) shows lower charging demand because only half of electric cars were charged at home during the night on a weekday as discussed in Section 3.6. As depicted in Fig. 8(d), three charging profiles were observed in the commercial sites. The maximum charging demand in Fig. 8(d) was also observed during the non-operational hours. The charging demand in the daytime on a weekday in winter is higher than that in the evening because the charging demand in the daytime is the sum of the other half of electric cars charged at the workplace and electric buses charged at bus stations in the daytime.

Fig. 8(e) and (f) illustrates the EV charging load profiles on a weekday in summer in the residential and commercial sites respectively. Compared with Fig. 8(a), lower charging demand during the night was also observed in the residential sites as shown in

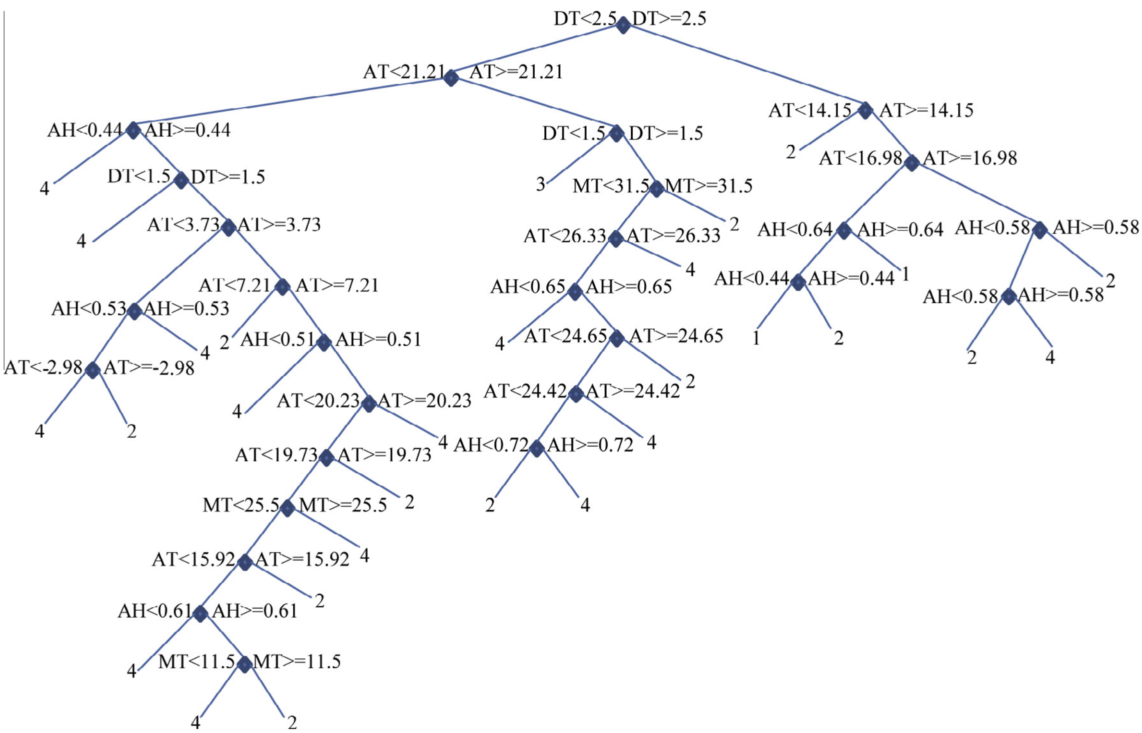

Fig. 9. Decision tree for cars.

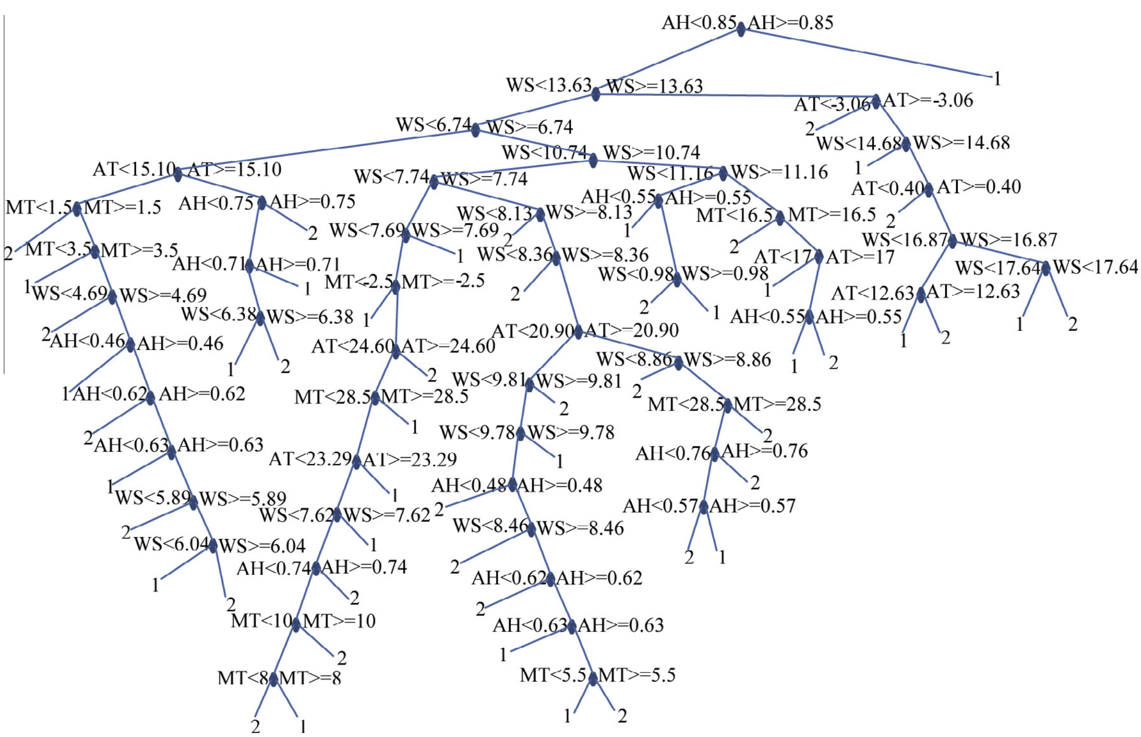

Fig. 10. Decision tree for buses.

Fig. 8(e) because only 50% of electric cars were charged at home on a weekday as aforementioned. As depicted in Fig. 8(f), three charging load profiles were observed in the commercial sites. The commercial EV charging demand in the daytime on a weekday in summer as shown in Fig. 8(f) accounts for the other half of electric cars charged at the workplace. The charging load profiles in the evening and in the non-operational hours in Fig. 8(f) account for the charging demand of electric buses charged at their bus stations or respective parking lots. As shown in Fig. 8(f), the maximum charging demand in the commercial sites resulted from charging buses during the non-operational hours.

Fig. 8(g) and (h) shows the EV charging demand on a weekend in summer in the residential and commercial sites respectively. It is worthy of notice that this fourth case study date (i.e., 8/15/2015) is a Korean holiday as well as a weekend. As shown in Fig. 8(g), the charging load profile in the residential areas shows high charging demand during the night because all cars were charged at home in the holiday that is also a weekend. Two charging demand profiles in Fig. 8(h) were observed in the commercial sites due to high traffic volume in the daytime as shown in Fig. 6(a). The maximum EV charging demand in the commercial sites during the non-operational hours was observed as shown in Fig. 8(h) because there was only 1 h charging starting time interval.

5. Conclusions

This paper proposed an EV charging demand forecasting model with big data technologies. The historical real-world traffic distribution and weather condition data processed by big data technologies were considered in the EV charging demand forecasting model. These big data handling processes include a cluster analysis to classify traffic patterns on each cluster, a relational analysis to identify influential factors affecting the traffic patterns, and a decision tree to establish classification criteria. The example case studies to anticipate EV charging demand in the residential and commercial sites on weekdays and weekends in winter and summer seasons were presented in Section 4 using the proposed EV charging demand forecasting model. High EV charging demand in the residential sites was observed during the night on weekends because all cars were charged at home on weekends as discussed in Section 3.6. In the commercial sites, high charging demand was observed during the non-operational hours due to its lesser charging starting time interval in this period than that in the other charging periods. The proposed EV charging demand forecasting model in this paper may help on the research on the impact of EV charging demand on the power system. In addition, the proposed EV charging demand forecasting model may allow utility operators to plan the operation and generation profiles in the future power systems by predicting the EV charging demand in the residential and commercial sites. The presented EV charging demand model can also contribute to deciding investment and operation plans for adaptive EV charging infrastructures depending on EV charging demand.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2014R1A1A1036384).

Appendix A

See Figs. 9 and 10.

References

[1] Trigg T, Telleen P. Global EV outlook: Understanding the electric vehicle landscape to 2020. France, Paris: OECD/IEA; 2013. <https://www.iea.org/publications/globalevoutlook_2013.pdf> [accessed 12.14.15].
[2] KEEI Korea energy demand outlook. KEEI. Ulsan, South Korea; 2015. <http://www.keei.re.kr/keei/download/outlook/OL1702e.pdf> [accessed 12.18.15].
[3] Brouwer AS, Kuramochi T, van den Broek M, Faaij A. Fulfilling the electricity demand of electric vehicles in the long term future: an evaluation of centralized and decentralized power supply systems. Appl Energy 2013;107:33–51. http://dx.doi.org/10.1016/j.apenergy.2013.02.005.
[4] Salah F, Ilg J, Flath C, Basse H, van Dinther C. Impact of electric vehicles on distribution substations: a Swiss case study. Appl Energy 2015;137:88–96. http://dx.doi.org/10.1016/j.apenergy.2014.09.091.
[5] Mu Y, Wu J, Jenkins N, Jia H, Wang C. A spatial-temporal model for grid impact analysis of plug-in electric vehicles. Appl Energy 2014;114:456–65. http://dx.doi.org/10.1016/j.apenergy.2013.10.006.
[6] Wang H, Song Q, Zhang L, Wen F, Huang J. Load characteristics of electric vehicles in charging and discharging states and impacts on distribution systems. In: Sustainable Power Generation and Supply (SUPERGEN) Int Conference, Hangzhou, China. p. 1–7. http://dx.doi.org/10.1049/cp.2012.1837.
[7] Park WJ, Song KB, Park JW. Impact of electric vehicle penetration-based charging demand on load profile. J Electr Eng Technol 2013;8(2):244–51. http://dx.doi.org/10.5370/JEET.2013.8.2.244.
[8] Neaimeh M, Wardle R, Jenkins A, Yi J, Hill G, Lyons P, et al. A probabilistic approach to combining smart meter and electric vehicle charging data to investigate distribution network impacts. Appl Energy 2015;157:688–98. http://dx.doi.org/10.1016/j.apenergy.2015.01.144.
[9] Leou RC, Su CL, Lu CN. Stochastic analyses of electric vehicle charging impacts on distribution networks. IEEE Trans Power Syst 2014;29(3):1055–63. http://dx.doi.org/10.1109/TPWRS.2013.2291556.
[10] Valsera-Naranjo E, Sumper A, Villafafila-Robles R, Martinez-Vicente D. Probabilistic method to assess the impact of charging of electric vehicles on distribution grids. Energies 2012;5:1503–31. http://dx.doi.org/10.3390/en5051503.
[11] Qian K, Zhou C, Allan M, Yuan Y. Modeling of load demand due to EV battery charging in distribution systems. IEEE Trans Power Syst 2011;26(2):802–10. http://dx.doi.org/10.1109/TPWRS.2010.2057456.
[12] Olivella-Rosell P, Villafafila-Robles R, Sumper A, Bergas-Jane J. Probabilistic agent-based model of electric vehicle charging demand to analyse the impact on distribution networks. Energies 2015;8:4160–87. http://dx.doi.org/10.3390/en8054160.
[13] Bae S, Kwasinski A. Spatial and temporal model of electric vehicle charging demand. IEEE Trans Smart Grid 2012;3(1):394–403. http://dx.doi.org/10.1109/TSG.2011.2159278.
[14] Liang H, Sharma I, Zhuang W, Bhattacharya K. Plug-in electric vehicle charging demand estimation based on queueing network analysis. In: Proc IEEE power and energy soc general meeting, National Harbor, MD. p. 1–5. http://dx.doi.org/10.1109/PESGM.2014.6939530.
[15] Lojowska A, Kurowicka D, Papaefthymiou G, van der Sluis L. From the transportation patterns to power demand: Stochastic modeling of uncontrolled domestic charging of electric vehicles. In: Proc. IEEE power and energy soc general meeting, San Diego, CA. p. 1–7. http://dx.doi.org/10.1109/PES.2011.6039187.
[16] Luo Z, Song Y, Hu Z, Xu Z, Yang X, Zhan K. Forecasting charging load of plug-in electric vehicles in China. In: Proc IEEE power and energy soc general meeting, San Diego, CA. p. 1–8. http://dx.doi.org/10.1109/PES.2011.6039317.
[17] Xydas E, Marmaras C, Cipcigan L, Hassan A, Jenkins N. Forecasting electric vehicle charging demand using support vector machines. In: Power engineering conference (UPEC), Dublin, Ireland. p. 1–6. http://dx.doi.org/10.1109/UPEC.2013.6714942.
[18] Dong X, Mu Y, Jia H, Wu J, Yu X. Planning of fast EV charging stations on a round freeway. IEEE Trans Sustain Energy 2016:1–10. http://dx.doi.org/10.1109/TSTE.2016.2547891. Early Access.
[19] Majidpour M, Qiu C, Chu P, Pota HR, Gadh R. Forecasting the EV charging load based on customer profile or station measurement? Appl Energy 2016;163:134–41. http://dx.doi.org/10.1016/j.apenergy.2015.10.184.
[20] Fiori C, Ahn K, Rakha H. Power-based electric vehicle energy consumption model: model development and validation. Appl Energy 2016;168:257–68. http://dx.doi.org/10.1016/j.apenergy.2016.01.097.
[21] Kelly J, MacDonald J, Keoleian G. Time-dependent plug-in hybrid electric vehicle charging based on national driving patterns and demographics. Appl Energy 2012;94:395–405. http://dx.doi.org/10.1016/j.apenergy.2012.02.001.
[22] Kara E, Macdonald J, Black D, Berges M, Hug G, Kiliccote S. Estimating the benefits of electric vehicle smart charging at non-residential locations: a data-driven approach. Appl Energy 2015;155:515–25. http://dx.doi.org/10.1016/j.apenergy.2015.05.072.
[23] Wang H, Zhang X, Ouyang M. Energy consumption of electric vehicles based on real-world driving patterns: a case study of Beijing. Appl Energy 2015;157:710–9.
[24] Xydas E, Marmaras C, Cipcigan L, Jenkins N, Carroll S, Barker M. A data-driven approach for characterising the charging demand of electric vehicles: a UK case study. Appl Energy 2016;162:763–71. http://dx.doi.org/10.1016/j.apenergy.2015.05.057.
[25] Li Y, Davis C, Lukszo Z, Weijnen M. Electric vehicle charging in China's power system: energy, economic and environmental trade-offs and policy implications. Appl Energy 2016;173:535–54. http://dx.doi.org/10.1016/j.apenergy.2016.04.040.
[26] Michalik P, Stofa J, Zolotova I. Concept definition for big data architecture in the education system. In: Proc of IEEE 12th int symp conference, Herl'any, Slovakia. p. 331–4. http://dx.doi.org/10.1109/SAMI.2014.6822433.
[27] Zhang P, Wu X, Wang X, Bi S. Short-term load forecasting based on big data technologies. CSEE J Power Energy Syst 2015;1(3):59–67. http://dx.doi.org/10.17775/CSEEJPES.2015.00036.
[28] MATLAB & Simulink. <http://www.mathworks.com/help/index.html> [accessed 02.20.16].
[29] Johnson R, Wichern D. Applied multivariate statistical analysis. 6th ed. New Jersey, USA: Pearson Education Inc.; 2007. p. 671–96.
[30] Saracli S, Dogan N, Dogan I. Comparison of hierarchical cluster analysis methods by cophenetic correlation. J Inequal Appl 2013;203:1–8. http://dx.doi.org/10.1186/1029-242X-2013-203.
[31] Sallehuddin R, Shamsuddin SM, Hashim SZM. Application of grey relational analysis for multivariate time series. In: 8th int conference on intelligent syst

design and applicat (ISDA), Kaohsiung, Taiwan. p. 432–7. http://dx.doi.org/10.1109/ISDA.2008.181.

[32] Park J, Park K. Overview of national electric vehicle deployment in Korea. Sejong, South Korea: KOTI; 2014. <http://2014excom.citynet-ap.org/wp-content/uploads/2014/12/5.4-KOTI-Transport-Session-EV-Deployment-in-Korea.pdf> [accessed 10.22.15].

[33] Korea Transportation Safety Authority <http://www.ts2020.kr/eng/main.do> [accessed 10.22.15].

[34] Cao Y, Tang S, Li C, Zhang P, Tan Y, Zhang Z, et al. An optimized EV charging model considering TOU price and SOC curve. IEEE Trans Smart Grid 2012;3(1):388–93. http://dx.doi.org/10.1109/TSG.2011.2159630.

[35] Society Automotive Engineers <http://standards.sae.org/j1772_201001/> [accessed 02.20.16].

[36] Zhang P, Qian K, Zhou C, Stewart B, Hepburn D. A methodology for optimization of power systems due to electric vehicle charging load. IEEE Trans Power Syst 2012;27(3):1628–36. http://dx.doi.org/10.1109/TPWRS.2012.2186595.

[37] Traffic Monitoring System. <http://www.road.re.kr/main/main.asp> [accessed 10.22.15].

[38] Hakimi S, Moghaddas-Tafrshi S. Impact of plug-in hybrid electric vehicles on Tehran's electricity distribution grid. In: World renewable energy congr, Linkoping, Sweden. p. 3589–96. http://dx.doi.org/10.3384/ecp110573589.

[39] Korean Statistical Information Service <http://kosis.kr/eng/> [accessed 10.22.15].