An Advanced Data Driven Model for Residential Electric Vehicle Charging Demand

Xiaochen Zhang
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, US
x.zhang@gatech.edu

Santiago Grijalva
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, US
sgrijalva@ece.gatech.edu

***Abstract*—As the electric vehicle (EV) is becoming a significant component of the loads, an accurate and valid model for the EV charging demand is the key to enable accurate load forecasting, demand respond, system planning, and several other important applications. We propose a data driven queuing model for residential EV charging demand by performing big data analytics on smart meter measurements. The data driven model captures the non-homogeneity and periodicity of the residential EV charging behavior through a self-service queue with a periodic and non-homogeneous Poisson arrival rate, an empirical distribution for charging duration and a finite calling population. Upon parameter estimation, we further validate the model by comparing the simulated data series with real measurements. The hypothesis test shows the proposed model accurately captures the charging behavior. We further acquire the long-run average steady state probabilities and simultaneous rate of the EV charging demand through simulation output analysis.**

***Index Terms*--Electric vehicles, load modeling, data mining, queuing analysis**

I. INTRODUCTION

Electric vehicles (EVs) draw and store energy from an electric grid to supply propulsive energy for the vehicle [1]. Since the US federal government highlighted electricity as a promising alternative to petroleum in the transportation sector in 2009 [2], the strong policy support has made US the leader of EV market. As of September 2014, the United States has the largest fleet of highway-capable EVs in the world, with about 260,000 plug-in electric cars sold since 2008 [3].

Many researchers have shown that in a high EV penetration environment, uncoordinated EV charging behavior could have a significant impact on distribution grids, especially on residential level [4-5]. Meanwhile, with a proper control strategy, the battery of the EV could potentially provide additional services to the grid through demand controls, such as flatting the peak load, providing voltage support and frequency regulation. In order to achieve these goals, it is crucial to develop an advanced model that captures the charging behavior of EVs for both operational and planning purposes.

Various research papers [6-10] model the EV charging as a queuing system. In reference [6], the EV charging time and duration are determined in a deterministic manner by some market signals and a fixed distance distribution. In reference [7], a $M/M/N_{\max}$ queue is introduced, where the EV arrives as a Poisson process with an exponentially distributed charging time, and N_{max} is the total charging capacity. Reference [8] employs an $M/M/\infty$ queue to capture the fact that residential EV charging is a self-service system. Both reference [7] and [8] assume that the EV charging arrival rate is not related to the number of EVs that are already in charging. The $M/M/s$ models in reference [9] and [10] are based on the assumption that the arrival process of EV charging event is a homogeneous Poisson process with a constant rate, and that the charging duration is exponentially distributed. Although we can derive the long-run average properties of the abovementioned models analytically, most of these models are based on some unrealistic assumptions without validation.

Thanks to the widely installed smart meters and corresponding infrastructure, for the first time, researchers and utilities have been able to gain access to the energy consumption patterns of consumers of a great resolution and at such a large scale [11]. In this paper, we propose a novel data driven approach to establish a valid model for residential EV charging demand by applying big data analytics on measurements directly collected from EV charging decks. Although EV charging behaviors is related to factors such as location, customer job, or even the gas price, the smart meter reading alone can be a good indicator which summarizes all these social factors. The proposed model allows us to capture the non-homogeneity and periodicity of the EV charging demand. Moreover, we estimate the EV charging duration with an empirical pdf generated from the real smart meter data.

The proposed new model does not require any of the pre-assumption mentioned above. And the model can be further

This work is supported by the Georgia Institute of Technology.

utilized by electric utilities for enhanced projection of EV demand and deployment of advanced coordination applications as part of demand response and grid services procurement.

The remainder of this paper is structured as follows. In Section 2, we derive the proposed model by removing unrealistic assumptions made by the simplest $M/M/\infty$ queue. In section 3, we estimate the parameters for the proposed model using real EV charging measurements and validate the model through a hypothesis test. In Section 4, we further illustrate the charging behavior of residential EVs and calculate the simultaneous rate through long-run average steady state statistics. And we conclude the paper in Section 5.

II. Modeling of EV Charging Demand

A. *Data observation*

The key advantage of the data driven EV model is that the model is supported by real smart meter measurements. The smart meter data not only provide us with the knowledge of residential EV charging patterns, but also plays a vital role in model validation.

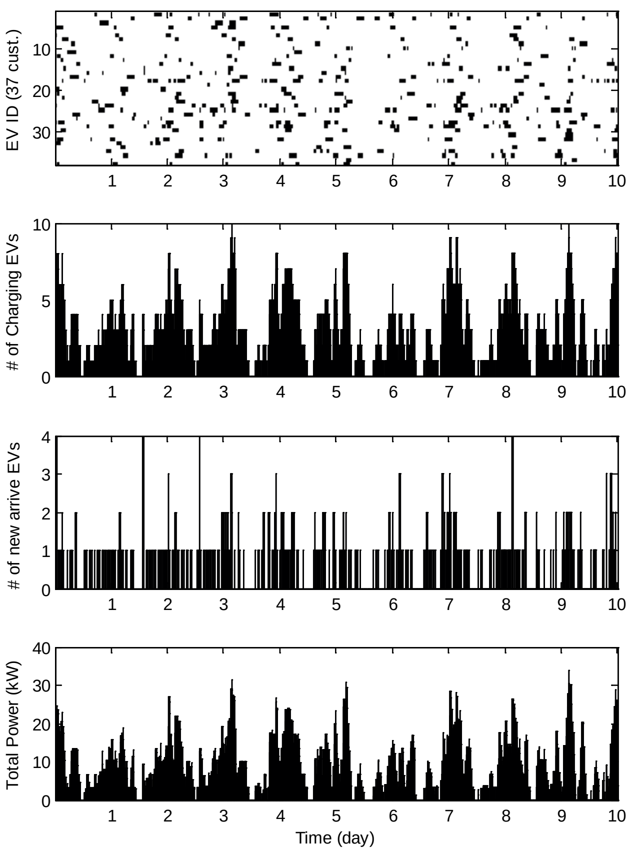

Figure 1. Observation of the EV charging behavior

Figure 1 is some general observations of 37 independent EVs behaviors collected by Pecan Street Inc. [12], Austin, Texas. The data were collected every 15 minutes directly from EV charging deck. In Figure 1.1, Black bars represent charging behaviors for the 37 EVs; Figure 1.2 shows the number of charging EVs through time; Figure 1.3 visualizes the number of EVs that start charging during each 15 minutes time interval; Figure 1.4 shows the energy consumption of all EVs. By observing the four plots, we claim the key of modeling EV charging demand (shown in Figure 1.4) is the modeling of Figure 1.2 through time, which can be further derived from EV charging duration (shown in Figure 1.1) and EV charging arrival rate (shown in Figure 1.3).

B. *General $M_1/M_2/\infty/N_{max}$ model*

The $M_1/M_2/\infty/N_{max}$ queue is the most widely adopted stochastic model for EV charging demand. In the model:

- M_1 means that the arrival of EV charging events follow a Poisson process with rate λ;
- M_2 means that the EV charging durations are independently and identically distributed (i.i.d.) with an exponential distribution of rate μ;
- ∞ refers to the infinite number of servers in the queuing system. In other words, the residential EV charging system is a self-service system with no waiting time;
- N_{max} refers to the total number of EVs in the community.

Let $X(t)$ be the number of charging EVs at time t, and the state space of $X(t)$ be S, where $S = \{1, 2, \dots, N_{max}\}$. Then, Figure 2 illustrates the transition diagram of the $M_1/M_2/\infty/N_{max}$ queuing system.

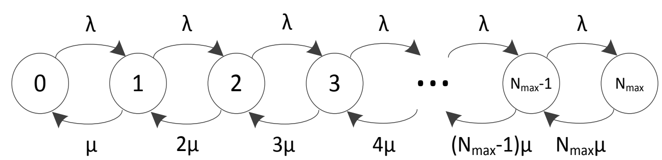

Figure 2. Transition diagram of $M_1/M_2/\infty/N_{max}$ queue

The advantage of using $M_1/M_2/\infty/N_{max}$ model lies in that researchers can derive the long-run average steady state probabilities of the system analytically. Let P_n denote the system's long-run average steady state probability of having n EVs charging simultaneously, then P_n can be given directly as

$$P_n = \frac{C_n}{e^{\lambda/\mu}}, \tag{1}$$

where $C_n = \frac{\lambda^n}{n!\mu^n}$ and $n = 1, 2, \dots, N_{max}$.

However, some pre-assumptions made by the $M_1/M_2/\infty/N_{max}$ model are not necessarily realistic, which requires further discussions.

C. *$M_1/M_2/\infty/N_{max}$ queue with finite calling population*

To begin with, the $M_1/M_2/\infty/N_{max}$ model assumes the arrival rate of new EV charging event remains the same no matter how many EVs are already in the charging state. However, this is not true as long as the number of EVs is finite. In a community with finite number of EVs, the potential new arrival rate of new EV charging event decreases as the number of charging EVs increases. In other words, let λ_i be the arrival rate when there are i EVs in the system, for any two integers $\{a, b: 0 \le a < b \le N_{max}\}$, we have $\lambda_a > \lambda_b$.

To model the finite number of residential EVs, we introduce the finite calling population model [13] for the $M_1/M_2/\infty/N_{max}$ queue. Assume each EV arrives independently according to a Poisson process with rate λ, then

$\lambda_i = (N_{max} - i)\lambda$. Figure 3 shows the transition diagram of the system with finite calling population.

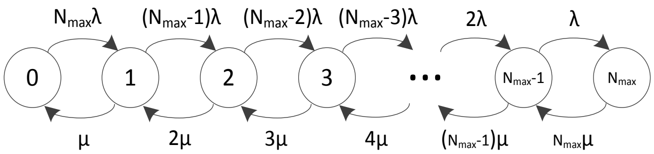

Figure 3. Transition diagram of the finite calling population model

Another advantage of adopting the finite calling population strategy is making the model scalable and more robust. Under the finite calling population strategy, instead of estimating the behavior of all N_{max} EVs, we estimate the behavior of every single EV. As long as the assumption that all EVs behavior independently holds, we could easily fit the model into systems with arbitrary number of EVs.

D. Non-homogeneous Poisson arrive rate

Another assumption made in $M_1/M_2/\infty/N_{max}$ model is that the arrival rate of EV charging events is a constant throughout the time. However, according to Figure 1.3, the arrival rate of EV charging events is not constant through time and has a period of 24 hours. Figure 4 shows the daily average EV charging arrival rate of the 37 residential EVs.

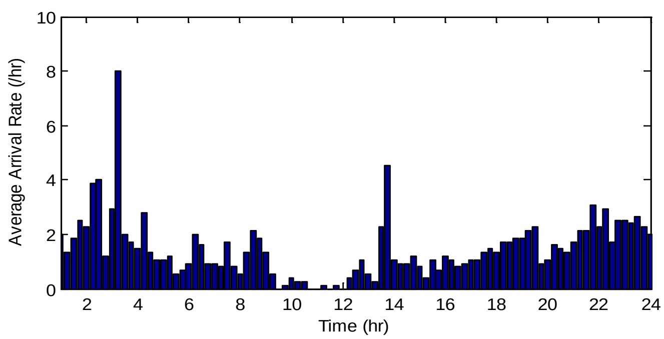

Figure 4. Av. daily arrival rate of nonhomogeneous Piosson model

To illustrate the periodicity of the arrival rate, Figure 5 shows the autocorrelation of the arrival rate with the lag resolution of every 15 minutes. Since the autocorrelation sequence has the same cyclic characteristics as the original arrival rate sequence, Figure 5 can serve to determine and verify the daily periodicity. As expected, the autocorrelation peaks in Figure 5 verify the daily periodicity of the arrival rate.

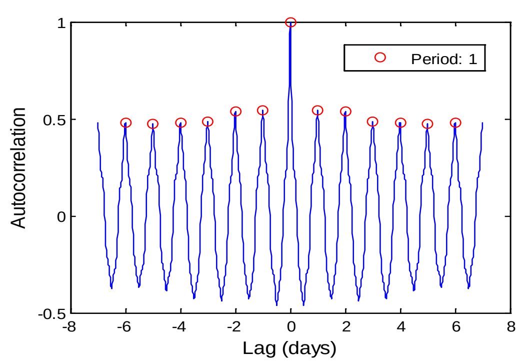

Figure 5. Lag autocorrelation plot of the arrival rate series (30 days)

To capture the time-variant property of EVs, we adopt a non-homogeneous Poisson process with a time dependent rate $\lambda(t)$. Let $m(t) = \int_0^t \lambda(t)dt$, according to the property of non-homogeneous Poisson process, the number of new arrivals from $t = t_0$ to $t = t_1$ follows the Poisson distribution of rate $\lambda = m(t_1) - m(t_0)$.

E. General $M_1/G/\infty/N_{max}$ model

Another assumption made by the $M_1/M_2/\infty/N_{max}$ model is that the charging duration of EVs is exponentially distributed. We will show this assumption is not valid through the memoryless property of the exponential distribution [14].

Assume an EV starts charging at time $t = 0$. Let $P(t > T)$ stand for the probability that the charging duration t is greater than T hours, and $P(t > T + S|t > S)$ the conditional probability of the charging more than $T + S$ hours given S hours of charging. According to the memoryless property of the exponential distribution, $P(t > T) = P(t > T + S|t > S)$. This contradicts to the common knowledge of EV charging behavior, since the battery capacity of EVs is limited.

To better model the EV charging duration, we adopt an empirical charging time distribution estimated from real EV charging measurements.

III. Model Estimation and Validation

As mentioned in the previous section, the data driven model developed in this paper is based on the historical data of 37 residential EVs for two months. One month of data are used for model training and parameter estimation (training data set), and the other month of data model validation (validation data set).

A. Model Parameter Estimation

According to Section II, we seek to model the residential EV charging behavior through a $M_t/G/\infty/N_{max}$ queue with a finite calling population, where M_t stands for the periodic non-homogeneous arrival rate; G stands for the empirical distribution of EV charging duration; ∞ means the charging system is a self-serve system with no waiting time; and N_{max} is the number of EVs in the community, which is known.

1) Estimation of the non-homogeneous arrival rate

Given the smart meter data resolution, we divide 24 hours of a day into 96 equal time intervals each with the length of Δt , then we treat the non-homogeneous arrival rate as piecewise constant in each time interval.

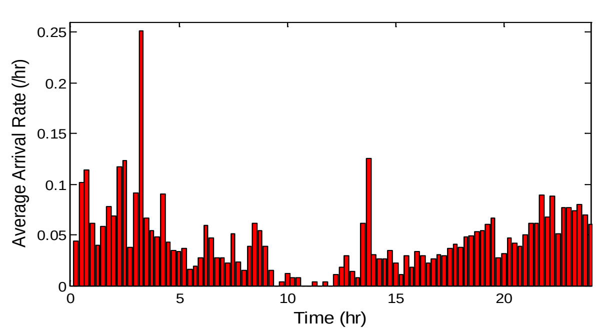

Figure 6. Estimated charging arrival rate per EV

Let $\lambda(k)$ be the arrival rate of each EV during time interval $((k-1)\Delta t, k\Delta t)$, where k is a discrete integer from 1 to 96. Let $W(k)$ and $N(k)$ be the number of existing and new arrivals of EVs during the time interval. Then $\lambda(k)$ can be estimated through

$$\hat{\lambda}(k) = \frac{N(k)/\Delta t}{N_{max} - W(k-1)}. \quad (2)$$

Figure 6 visualizes the daily average arrival rate for each EV through time using one month of training data.

2) Estimation of EV charging duration

Instead of using exponential distribution, we capture the EV charging duration through an empirical distribution observed from the training data set. Figure 7 shows the empirical probability density function (pdf) of the EV charging duration.

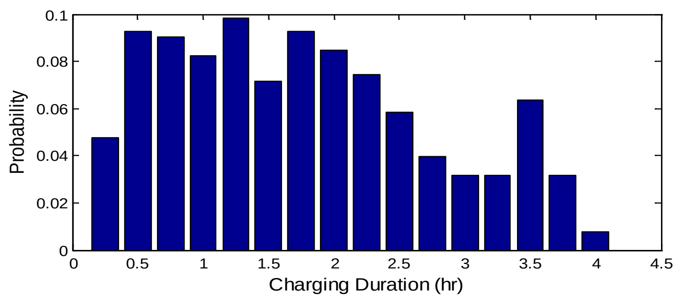

Figure 7. The emperical pdf of the EV charging duration

B. Model Validation

Upon the establishment of the model, we further validate it by comparing the simulated data with the validation data.

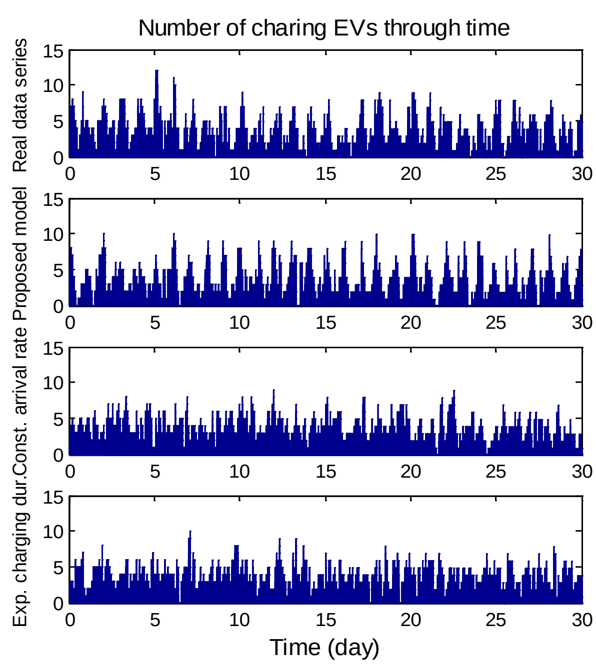

Figure 8. Comparison between simulated and validation data series

Figure 8 compares our model with real measurements and two other widely used queuing models. From Figure 8.3, we see that if we model the arrival rate as a constant through time, we lose the periodicity and the time variant property of the real measurements. From Figure 8.4, we can see that adopting an exponentially distributed charging duration will distort the true charging behaviors by having charging durations longer than 4 hours, which is unlikely to happen [15]. From Figure 8.2, the simulated data series generated by our model is stable and behavior very similar to the real measurements in Figure 8.1. To validate the model analytically, we run the simulation 100 times (100 replications) each with the length of 100 days. In each replication, the first 10 days' data are trimmed to ensure the data stability.

Let $\bar{D}_k$ be the average number of charging EVs during the kth time interval estimated using the validation data, where $k = 1, 2, \ldots, 96$. Similarly, let $\widehat{D}_{k,i}$ be the average number of charging EVs during the same time interval estimated by the ith replication. To this end, for each replication, define the difference $G_i = \widehat{D}_{k,i} - \bar{D}_k$, where $i = 1, 2, \ldots, 100$.

If the proposed model captures the true EV charging behavior well, G_i should be approximately normally distributed with mean $\mu_g = 0$ and variance σ_g^2 [16]. As a result, we construct a hypothesis test where,

$$\begin{cases} H_0: \mu_g = 0 \\ H_1: \mu_g \neq 0 \end{cases}. \quad (3)$$

Under the null hypothesis, the statistic

$$t_{N_2-1} = \frac{\bar{G} - \mu_g}{S_g/\sqrt{N_2}}, \quad (4)$$

follows the t distribution with $N_2 - 1$ degrees, where N_2 is the number of the replications, $\bar{G}$ and S_g are sample mean and sample variance [16]. TABLE I compares the statistics $\bar{G}$ and S_g corresponding to the three above mentioned models. It is clear that the proposed model has smaller mean and variance, which means it's a better model of the real EV charging behaviors.

TABLE I. MODEL COMPARISON

Model Type	sample mean $\bar{G}$	sample variance S_g
Constant Arrival Rate Model	-0.0186	1.6199
Constant Charging Rate Mode	-0.0558	0.6345
Proposed Model	-0.0064	0.5109

Given the significance level of $\alpha = 0.05$, we compute the confidence interval for μ_g, which is $(-0.1455, 0.1991)$. Since the interval contains zero, we cannot reject H_0 at the given significance level, which validates the proposed model as a good representation of the EV charging behavior.

IV. OUTPUT ANALYSIS

In order to obtain the long-run average steady state property of the proposed EV charging model, we set the simulation replications to 100, and each replication with the length of 100 days. Similarly we curtails the first 10 days of each replication due to stability requirements.

A. Long-run average number of charging EVs

Figure 9 shows the long-run average number of charging EVs throughout a day (blue curve). The 25^{th} and 75^{th} percentiles are also drawn respectively (red and green curves). All three curves suggest that the residential EV charging peak occurs during the night and that the span between 25^{th} and 75^{th}

percentiles are relatively small compared to the total EV number of 37.

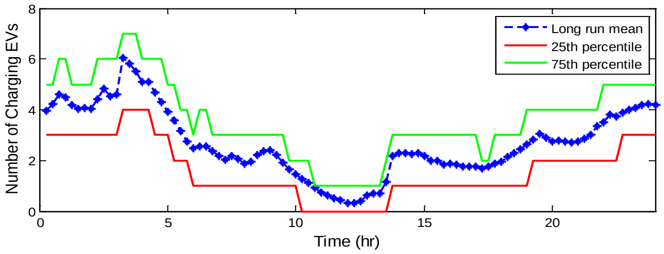

Figure 9. The long-run average, 25th and 75th percentile curves

B. Long-run average steady state probabilities

Let a $N_{max} \times 96$ matrix $\boldsymbol{P}$ be the long-run average steady state probability matrix, where $\boldsymbol{P}(n,k)$ denotes the long-run steady state probability of having n EVs charging during time interval k, then for each $k = 1,2,\dots,96$, we have

$$\sum_{n=1}^{N_{max}} \boldsymbol{P}(n,k) = 1. \qquad (5)$$

We visualize the long-run probabilities of the system through Figure 10, where the color in the plot represents the possibility of have n EV charging at a given time t.

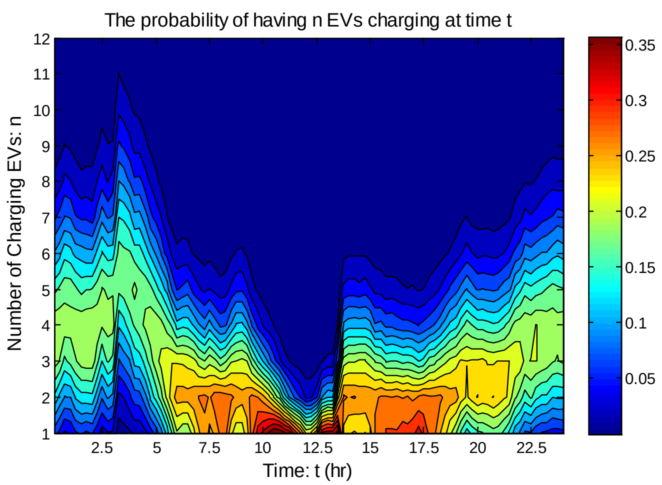

Figure 10. Visualization of the P matrix

Let ρ be the simultaneous rate of the EV charging load. Define ρ as $\rho = \text{Charging EV number}/\text{Total EV number}$ during the peak EV charging time. Then, the cumulative density function of ρ, which is $P(\rho \le \rho_0)$, provides essential information to estimate the simultaneous rate of EVs. For example, from matrix $\boldsymbol{P}$, we have $P(\rho \le 12/37) \ge 98.5\%$. This implies that for a community with 37 EVs, even in the worst case, the possibility of having 12 or more EVs charging simultaneously during one day is very slim (less than 1.5%).

V. CONCLUSION

This paper proposes a novel data driven model for residential EV charging demand. Compared with other queuing models, the proposed model allows us to capture the non-homogeneity and periodicity of the EV charging demand, and to estimate the charging duration with an empirical pdf. Upon parameter estimation, we validate the model through hypothesis testing and further acquire the EV charging long-run average probabilities and simultaneous rate through simulation output analysis. The proposed method can be utilized by electric utilities for enhanced projection of EV demand and deployment of advanced coordination applications as part of demand response and grid services procurement.

Further studies may include the analytical deriving of the long-run average steady state statistics for the EV charging behavior and the development of corresponding demand respond control based on the proposed EV load model.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of Dr. Sigrún Andradóttir to the model formulation and Dr. Jim Shead to the data acquisition.

REFERENCES

[1] Bradley, T. H. and A. A. Frank, "Design, demonstrations and sustainability impact assessments for plug-in hybrid electric vehicles," *Renewable and Sustainable Energy Reviews*, vol. 13(1), pp. 115-128, 2009.

[2] Carley, S., et al., "Intent to purchase a plug-in electric vehicle: A survey of early impressions in large US cites," *Transportation Research Part D: Transport and Environment*, vol. 18(0), pp. 39-45, 2013.

[3] Jeff Cobb, "Global Plug-in Car Sales Now Over 600,000," *HybridCars.com*, Retrieved 2014-10-23.

[4] Clement-Nyns, K., et al., "The Impact of Charging Plug-In Hybrid Electric Vehicles on a Residential Distribution Grid," *IEEE Trans. Power Systems*, vol. 25(1), pp. 371-380, 2010.

[5] Sortomme, E., et al. "Coordinated Charging of Plug-In Hybrid Electric Vehicles to Minimize Distribution System Losses," *IEEE Trans. Smart Grid*, vol. 2(1), pp. 198-205, 2011.

[6] Qian, Kejun, et al., "Modeling of load demand due to EV battery charging in distribution systems," *IEEE Trans. Power Systems*, vol. 26(2), pp. 802-810, 2011.

[7] Gan, L. and Z. Xiao-Ping, "Modeling of Plug-in Hybrid Electric Vehicle Charging Demand in Probabilistic Power Flow Calculations," *IEEE Trans. Smart Grid*, vol. 3(1), pp. 492-499, 2012.

[8] Peng, H., et al., "A Novel Coordinative Resident Electric Vehicle Charging Mechanism," in *2012, Power and Energy Engineering Conf.*

[9] Said, D., et al., "Queuing model for EVs charging at public supply stations," in *2013, Wireless Communications and Mobile Computing Conf.*, vol. 1(5), pp. 65-70

[10] Turitsyn, K., et al., "Robust Broadcast-Communication Control of Electric Vehicle Charging," in *Proc. 2010, IEEE Smart Grid Communications International Conf.*, vol. 4(6), pp. 203-207.

[11] Zhang, X., et al., "A time-variant load model based on smart meter data mining," in *Proc. 2014 IEEE Power Engineering Society General Meeting*, vol. 1(5), pp. 27-31.

[12] Pecan Street Dataport - A Universe of Data, Available Around the World. Available: https://dataport.pecanstreet.org/.

[13] D. Anderson, "An Introduction to Management Science: Quantitative Approaches to Decision Making," 13nd ed., Cengage Learning, 2011, pp. 256.

[14] Feller, W., "Introduction to Probability Theory and Its Applications," vol. 2, 2nd ed., Wiley, 1971, Section 1.3.

[15] Yijia, C., et al., "An Optimized EV Charging Model Considering TOU Price and SOC Curve," *IEEE Trans. Smart Grid* vol. 3(1), pp. 388-393, 2012.

[16] Altiok, T., et al., "Simulation Modeling and Analysis with ARENA," Elsevier, 2007, pp. 161.