

An Advanced Data Driven Model for Residential Plug-in Hybrid Electric Vehicle Charging Demand

Xiaochen Zhang

Georgia Institute of Technology
School of Electrical and Computer Engineering
Advanced Computational Electricity Systems Laboratory
Atlanta, GA
x.zhang@gatech.edu

Santiago Grijalva

Georgia Institute of Technology
School of Electrical and Computer Engineering
Advanced Computational Electricity Systems Laboratory Line
Atlanta, GA
sgrijalva@ece.gatech.edu

Abstract—As the plug-in hybrid electric vehicle (PHEV) is becoming a very significant component in residential loads, an accurate and valid model for the PHEV charging demand is the key for load forecast, demand respond, system planning and so forth. As a result, we propose a data driven queuing model for residential PHEV charging demand by performing data analytics on smart meter measurements. The data driven model captures the non-homogeneity and periodicity of the residential PHEV charging behaviors 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 with real measurements. The hypothesis test shows the proposed model captures the charging behaviors well. And we acquire the long-run steady state performance of the PHEV charging demand through simulation output analysis.

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

I. INTRODUCTION

Plug-in hybrid electric vehicles (PHEVs) are electric vehicles that can 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 PHEV market. As of June 2014, the monthly delivered PHEV are over 126,000. The PHEV is becoming a very significant load component in the US.

Many researchers have shown that in a high PHEV penetration environment, uncoordinated PHEV charging behaviors could have a significant impact on distribution grids, especially on residential level. [3-4] Meanwhile, with a proper control strategy, the battery of the PHEV could potentially provide additional services to the grid through demand controls, such as flattening the peak load, providing voltage support and so forth. However, to achieve these goals, it is crucial to develop

an advance model to capture the charging behaviors of PHEVs for both operational and planning purposes.

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 in great resolution at such a large scale. [5] In this paper, we propose a novel data driven approach to establish a valid model for residential PHEV charging demand by applying the big data analytics on real PHEV charging measurements.

Many literatures model the PHEV charging behaviors as a queuing system. First, to get the long-run average steady state results analytically, most of these models are based on some unrealistic assumptions without validation. Second, most of the proposed models are not validated through real PHEV charging measurements, which could lead to inaccurate representations of the real charging behaviors. In reference [6], the PHEV charging duration and the starting charging time are determined by a fixed distance distribution and market signals respectively in a deterministic manner. In reference [7], a $M/M/s/N_{max}$ model is introduced, where N_{max} is the total capacity for PHEV charging. Reference [8] employs an $M/M/\infty$ queue to capture the fact that residential PHEV charging is a self-service system. Both reference [7] and [8] assume that the arrival rate of the PHEV charging is not related to the number of PHEVs that are already in charging. Reference [9] and [10] adopt an $M/M/s$ queuing model, based on the assumption that the arrival process of PHEV charging event is a homogeneous Poisson process with constant rate, and that the charging duration is exponentially distributed.

In this paper, we propose a nonhomogeneous queuing model with a finite calling population to capture the charging behaviors of residential PHEVs. The proposed new model does not require any of the pre-assumption mentioned above thanks to the smart meter database. We estimate and validate the proposed model with real PHEV charging data collected by the Pecan Street Project. [11]

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 $M/M/\infty$ queuing model. In section 3, we estimate the parameters for the proposed model using real PHEV charging measurements and validate the model through hypothesis testing. In Section 4, we perform output analysis and further illustrate the charging behaviors of residential PHEVs through long-run average steady state statistics. Finally, we conclude the paper in section 5.

II. MODELING OF PHEV CHARGING DEMAND

A. Data observation

The key advantage of a data driven PHEV model is that the model is backed up by real smart meter measurements. The smart meter data not only provide us with knowledge of residential PHEV charging patterns to build a model and perform parameter estimation, but also play a vital role in model validation.

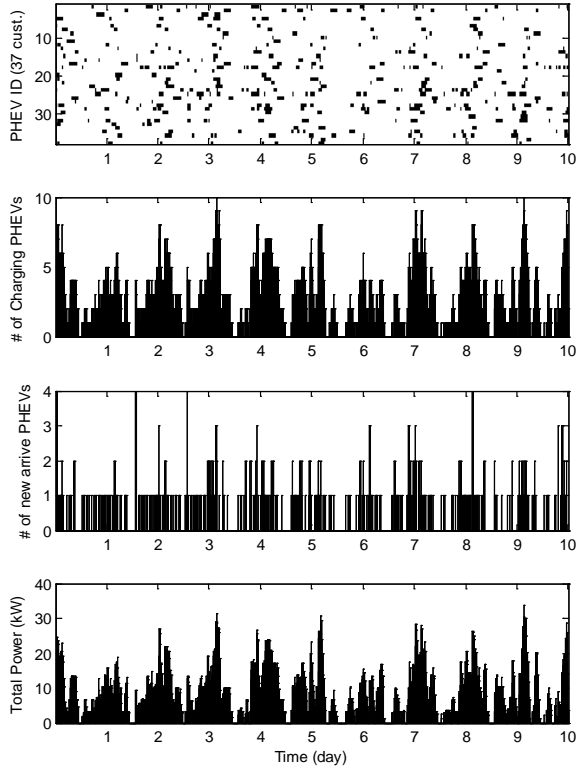


Figure 1. Observation of the PHEV charging behaviors

Figure 1 provides some general observations of 37 independent PHEVs for 10 days, where the PHEV charging behaviors are monitored by separate smart meters. The data were collected every 15 minutes by Pecan Street Inc., Austin, Texas. Black bars in Figure 1.1 represent the PHEV charging for the 37 customers through time; Figure 1.2 shows the number of charging PHEVs through time; Figure 1.3 visualizes the number of PHEVs that start charging during each 15 minutes time interval; Figure 1.4 shows the energy consumption of all PHEVs. By observing the four plots, we claim the key of modeling PHEV charging demand (see Figure 1.4) is the modeling of Figure 1.2 through time, which can be further

derived from PHEV charging duration (see Figure 1.1) and PHEV charging arrival rate (see Figure 1.3).

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

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

- M_1 stands for that the arrival of PHEV charging events follow a Poisson process with rate λ ;
- M_2 stands for that the charging duration of PHEVs is independently and identically distributed with an exponential distribution with rate μ ;
- ∞ refers to the infinite number of servers in the queuing system. In other words, the residential PHEV charging system is a self-service system with no waiting time;
- N_{max} refers to the total number of PHEVs in the community.

Let $X(t)$ be the number of PHEVs in charging state 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.

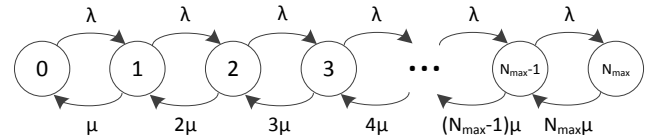


Figure 2. Transition diagram of $M_1/M_2/\infty/N_{max}$ queue

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 analytically. Let P_n denote the system's long-run average steady state probability of having n PHEVs in charging simultaneously, then P_n can be given directly as

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

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 discussion.

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 PHEV charging event remains the same no matter how many PHEVs are already in charging state. In fact, this is not true as long as the number of PHEVs is finite. In a community with finite number of PHEVs, the potential new arrival rate decreases as the number of PHEVs in charging state increases. In other words, let λ_i be the arrival rate when there are i PHEVs in the system, for any two integers $\{a, b: 0 \leq a < b \leq N_{max}\}$, we have $\lambda_a > \lambda_b$.

To model the finite residential PHEV number, we introduce the finite calling population model [12] for the $M_1/M_2/\infty/N_{max}$ queue. Assume each PHEV 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.

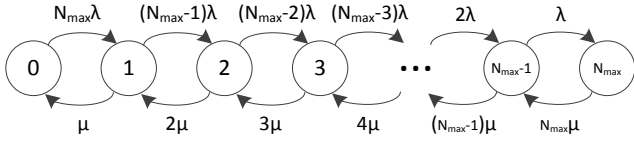


Figure 3. Transition diagram of the finite calling population model

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 behaviors of all N_{max} PHEVs, we estimate the behaviors of every single PHEV. As long as the assumption that all PHEVs behavior independently holds, we could easily fit the model into systems with arbitrary number of PHEVs.

D. Non-homogeneous Poisson arrive rate

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

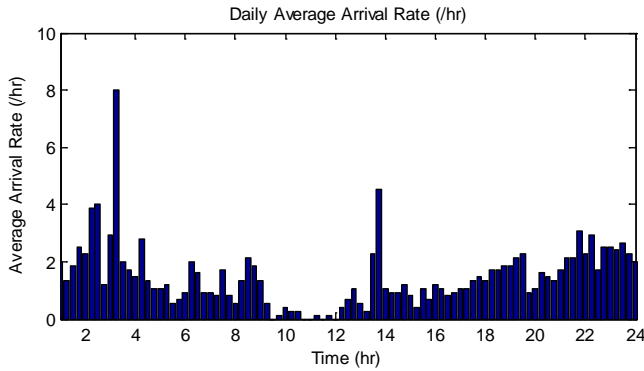


Figure 4. Average daily arrival rate of nonhomogeneous Piosson model

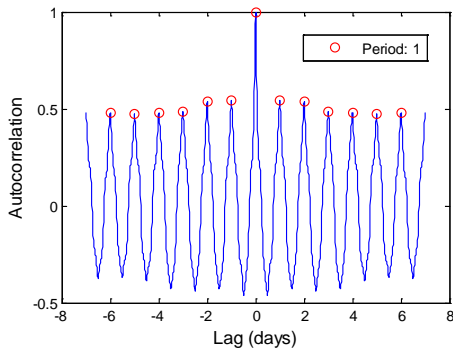


Figure 5. Lag autocorrelation plot of the arrival rate series (30 days)

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.

To capture the time-variant property of PHEVs, 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 PHEVs is exponentially distributed. We will show this assumption is not valid through the memory less property of exponential distribution.

Assume a PHEV starts charging at time $t = 0$. Let $P(t > T)$ stand for the probability of the charging duration is greater than T hours, and $P(t > T + S | t > S)$ the probability of the charging duration is greater than T hours after S hours of charging. According to the memory less property of exponential distribution, $P(t > T) = P(t > T + S | t > S)$. This contradicts to the common knowledge of PHEV charging behaviors, since the capacity of the PHEV battery is limited.

To improve the PHEV charging duration model, we adopt an empirical charging time distribution estimated from real PHEV 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 PHEVs 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 PHEV charging behaviors through a $M_t/G/\infty/N_{max}$ queue with finite calling population, where M_t means the periodic non-homogeneous arrival rate is a function of time t ; G stands for the empirical distribution of PHEV charging duration; ∞ means the charging system is a self-serve system with no waiting time; and N_{max} is the number of PHEVs 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 Δt , then we treat the non-homogeneous arrival rate as piecewise constant in each time interval.

Let $\lambda(k)$ be the arrival rate of each PHEV 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 PHEVs during the time interval. Then $\lambda(k)$ can be estimated through

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

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

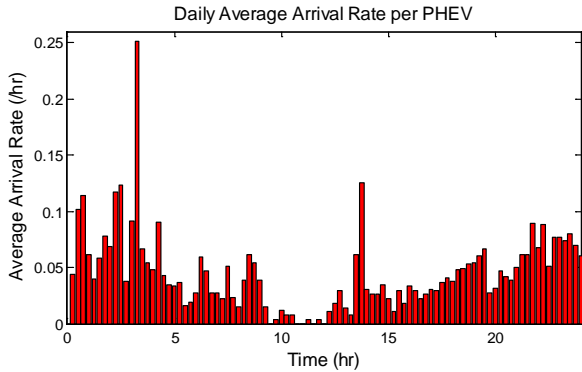


Figure 6. Estimated arrival charging rate per PHEV

2) Estimation of PHEV charging duration

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

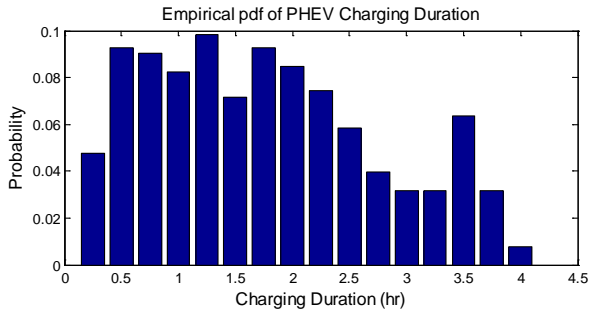


Figure 7. The empirical *pdf* of the PHEV charging duration

B. Model Validation

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

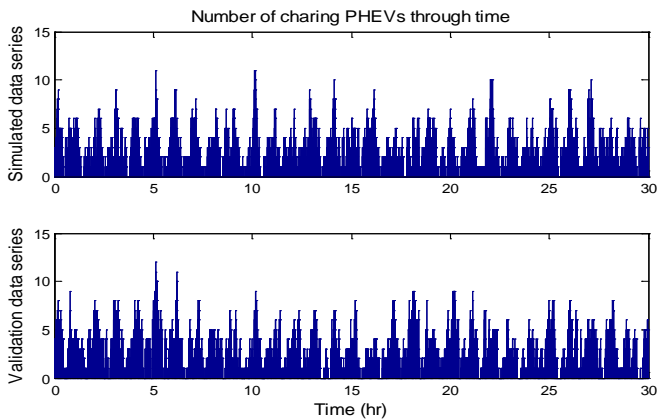


Figure 8. Comparison between simulated and validation data series

Figure 8 shows that the simulated data series is stable and behavior very similar to the real measurements in the validation data set. To validate the model analytically, we run the

simulation 100 times (100 replications) with the length of 100 days for each replication. 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 PHEVs during the k th time interval using the validation data, where $k = 1, 2, \dots, 96$. Similarly, let $\hat{D}_{k,i}$ be the average number of charging PHEVs during the same time interval in the i th replication of the simulation. To this end, for each replication, define the difference $G_i = \hat{D}_{k,i} - \bar{D}_k$, where $i = 1, 2, \dots, 100$.

If the model captures the PHEV charging behavior well, G_i should be approximately normally distributed with mean $\mu_g = 0$ and variance σ_g^2 [13]. As a result, we construct a hypothesis test where,

$$\begin{cases} H_0: \mu_g = 0 \\ H_1: \mu_g \neq 0 \end{cases}$$

Under the null hypothesis, the statistic

$$t_{N_2-1} = \frac{\bar{G} - \mu_g}{S_g / \sqrt{N_2}}$$

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.

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 PHEV charging behaviors.

IV. OUTPUT ANALYSIS

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

A. Long-run average number of charging PHEVs

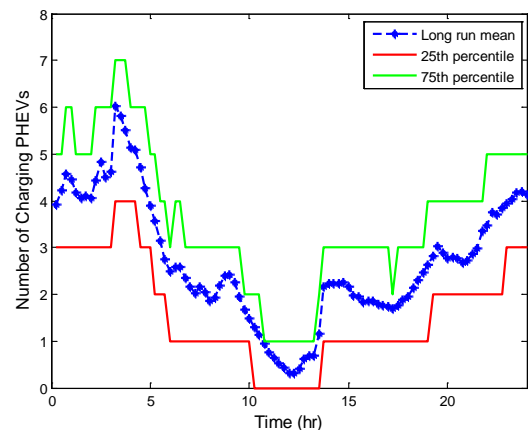


Figure 9. The long-run mean, 25th percentiles and 75th percentile plot

Figure 9 shows the long-run average number of charging PHEVs throughout a day (blue curve). The 25th and 75th percentiles of PHEV numbers are also drawn respectively (red and green curves). All three curves suggest that the residential

PHEV charging peak occurs during the night and that the span between 25th and 75th percentiles are relatively small compared to the total PHEV number of 37.

Let $W_{i,j,k}$ denote the number of charging PHEVs in the i th replication, on j th day, during the time interval k . Then, for each replication i ,

$$\bar{W}_{i,k} = \frac{1}{N_1} \sum_{j=1}^{N_1} W_{i,j,k}$$

is an unbiased estimator for the long-run average number of charging PHEVs in the k th time interval W_k , where N_1 is the length of each simulation replication after curtail. Similarly, the unbiased estimator for W_k using all N_2 replications is

$$\bar{W}_k = \frac{1}{N_2} \sum_{i=1}^{N_2} \bar{W}_{i,k}.$$

According to the central limit theorem, $\bar{W}_{i,k}$ is approximately normally distributed with mean $\mu_{W,k}$ and variance $\sigma_{\bar{W},k}^2$. As a result, we could construct the confidence interval for W_k using all N_2 replications of the simulated data at the significance level α . Let statistic

$$t_{N_2-1} = \frac{\bar{W}_k - \mu_{W,k}}{S_{W,k}/\sqrt{N_2}}.$$

Since statistic t_{N_2-1} follows the t distribution with $N_2 - 1$ degrees, we compute the confidence interval for W_k through

$$\bar{W}_k \pm t_{N_2-1, 1-\alpha/2} S_{W,k}/\sqrt{N_2},$$

where $k = 1, 2, \dots, 96$.

B. Long-run average steady state probabilities

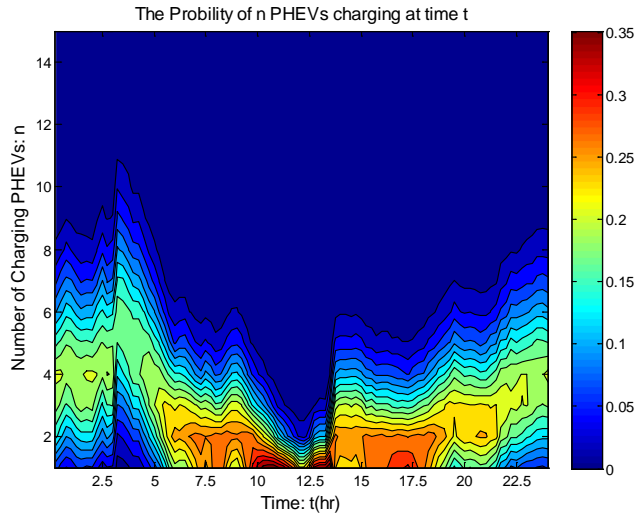


Figure 10. Visualization of the P matrix

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

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

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

V. CONCLUSION

This paper proposes a novel data driven model for residential PHEV charging demand. Compared with other queuing models, the proposed data driven model allows us to capture the non-homogeneity of the PHEV charging behaviors and the charging duration by taking advantage of the real measurements. Upon parameter estimation, we further validate the model through hypothesis testing. We acquire the long-run average PHEV charging behaviors and steady state probabilities through simulation output analysis.

Further studies may include the analytical deriving of the long-run average steady state statistics for PHEV charging behaviors and the development of corresponding demand respond control based on the proposed PHEV 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 (2009). "Design, demonstrations and sustainability impact assessments for plug-in hybrid electric vehicles." *Renewable and Sustainable Energy Reviews* 13(1): 115-128.
- [2] Carley, S., et al. (2013). "Intent to purchase a plug-in electric vehicle: A survey of early impressions in large US cities." *Transportation Research Part D: Transport and Environment* 18(0): 39-45.
- [3] Clement-Nyns, K., et al. (2010). "The Impact of Charging Plug-In Hybrid Electric Vehicles on a Residential Distribution Grid." *Power Systems, IEEE Transactions on* 25(1): 371-380.
- [4] Sortomme, E., et al. (2011). "Coordinated Charging of Plug-In Hybrid Electric Vehicles to Minimize Distribution System Losses." *Smart Grid, IEEE Transactions on* 2(1): 198-205.
- [5] Zhang, X., et al. (2014). A time-variant load model based on smart meter data mining. *PES General Meeting | Conference & Exposition, 2014 IEEE*.
- [6] Qian, Kejun, et al. "Modeling of load demand due to EV battery charging in distribution systems." *Power Systems, IEEE Transactions on* 26.2 (2011): 802-810.
- [7] Gan, L. and Z. Xiao-Ping (2012). "Modeling of Plug-in Hybrid Electric Vehicle Charging Demand in Probabilistic Power Flow Calculations." *Smart Grid, IEEE Transactions on* 3(1): 492-499.
- [8] Peng, H., et al. (2012). A Novel Coordinative Resident Electric Vehicle Charging Mechanism. *Power and Energy Engineering Conference (APPEEC), 2012 Asia-Pacific*.
- [9] Said, D., et al. (2013). Queuing model for EVs charging at public supply stations. *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*.
- [10] Turitsyn, K., et al. (2010). Robust Broadcast-Communication Control of Electric Vehicle Charging. *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*.
- [11] Pecan Street Dataport - A Universe of Data, Available Around the World. <https://dataport.pecanstreet.org/>
- [12] D. Anderson, " An Introduction to Management Science: Quantitative Approaches to Decision Making," 13nd ed., Cengage Learning, 2011, pp. 256.
- [13] Altiock, T., et al. " Simulation Modeling and Analysis with ARENA," Elsevier, 2007, pp. 161.

