Data Analytics of Real-World PV/Battery Systems

Miao Zhang, Zhixin Miao, Lingling Fan
Department of Electrical Engineering, University of South Florida

***Abstract*—This paper presents data analytic results based on four-year data from real-world 1.6 kW photovoltaic (PV) panels and 20 kWh Lithium-ion batteries installed at St. Petersburg Florida. The 1-minute interval raw data are collected and stored in spreadsheets. We present the raw data related to power outputs from PVs and batteries as well as estimated state-of-charge (SOC) of batteries. Data analysis is conducted using Python sqlite3 and Pandas to examine histograms of PV daily energy output and battery degradation.**

***Index Terms*—PV, battery, data analysis**

I. Introduction

Two photovoltaic-battery systems were installed at University of South Florida (USF) St. Petersburg campus (Campus Battery) and at Albert Whitted Airport at St. Petersburg downtown (Airport Battery) to realize smart grid functionalities such as peak shaving or demand response. Currently, each PV is connected to the grid through an inverter, while the two batteries are 5kW-4 hours Li-ion batteries and equipped with a charger and an inverter. Each battery has 16 battery cells. Each cell has a rated dc voltage 3 V and rated current 400 A. The rated dc voltage of each battery is 48 V. The ac side of the battery at the USF St. Petersburg campus is connected to a 120/208 V panel. The ac side of the battery at the Albert Whitted Airport is connected to a 120/240 V panel.

The configuration of the PV-battery system is shown in Fig. 1. The two batteries are operated in two modes. The first one is operated for peak shaving and energy shift. The second one is operated to realize demand response.

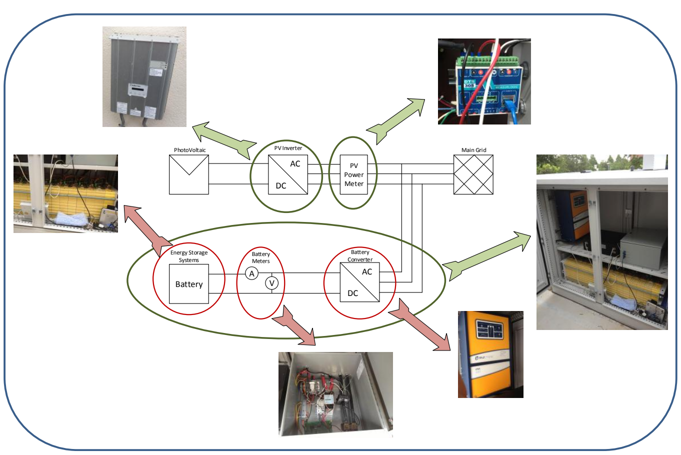


Fig. 1: Configuration of Photovoltaic-Battery systems in both campus and airport sites.

This project is supported by Duke Energy Florida through Sustainable Electrical Energy Delivery Systems.

(1) Peak shaving provided by a PV/battery system with constant output power. The PV/battery system is expected to provide constant output power at peak periods, Summer (14:00-20:00) and Winter (06:00-10:00). The net output of the SEEDS system (PV and battery) will be held at 1.4 kW. The battery will be charged to a minimum available energy of 10kWh prior to 6 am daily. The charging will commence at midnight and be done by 5 am daily. Off-peak energy and/or available solar PV energy will be used for the charging.

(2) Demand response by a PV/battery system with maximum output power. The second PV/battery system will also be charged during the off-peak period. Full 5 kW discharge capacity of the charged battery system and PV output will be delivered to the system whenever there is a command.

Approach and requirements to realize smart grid functions: Remote real-time control and monitoring system are required to develop the above mentioned smart grid functions. In order to realize the remote control and monitoring, the following requirements must be met:

(1) Measurements such as power, voltage, current flowing into or out from the ac side of the battery system should be obtained constantly. Energy can be computed based on these measurements.

(2) Measurements such as temperature, dc voltage, dc currents, battery SOC for a battery should be monitored.

(3) The human machine interfaces (HMI) provided by the battery vendor (Green Smith) should be able to execute inverter control to charge and discharge the battery system. At the SCADA control center, the USF personnel set the PV/battery operation patterns, including power dispatch level at every hour. At each PV/battery system site, the battery's controller receives this command and reads PV's power. It then sets its power demand to be the total power subtracted by the PV power.

The battery vendor Green Smith provide real-time measurements from both ac side and dc side measurements and battery SOC estimation. In this paper, we use SOC provides by the vendor. Our work on the battery system identification can be found at [1]. The battery SOC was estimated by AutoRegressive eXogenous (ARX) model, a technique that has been used for dynamic system parameter estimation for synchronous generators in our other previous work [2], [3].

The main contributions in this paper can be concluded as follows. (i) The approaches of storing and analysis of real-world big data using database and python can handle

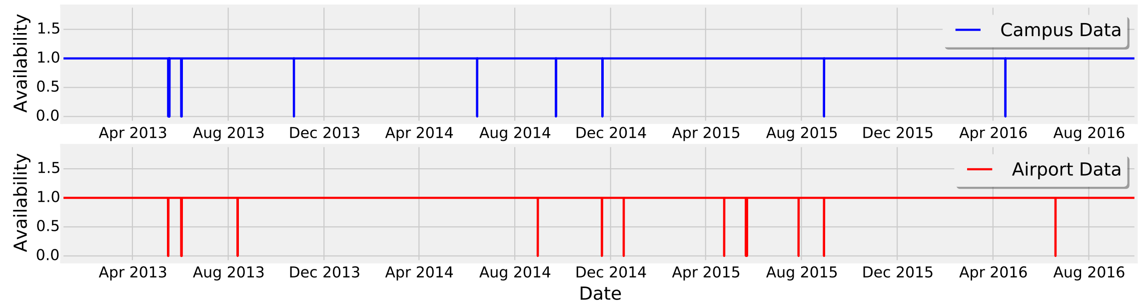


Fig. 2: 2013-2016 Campus and Airport Data Availability.

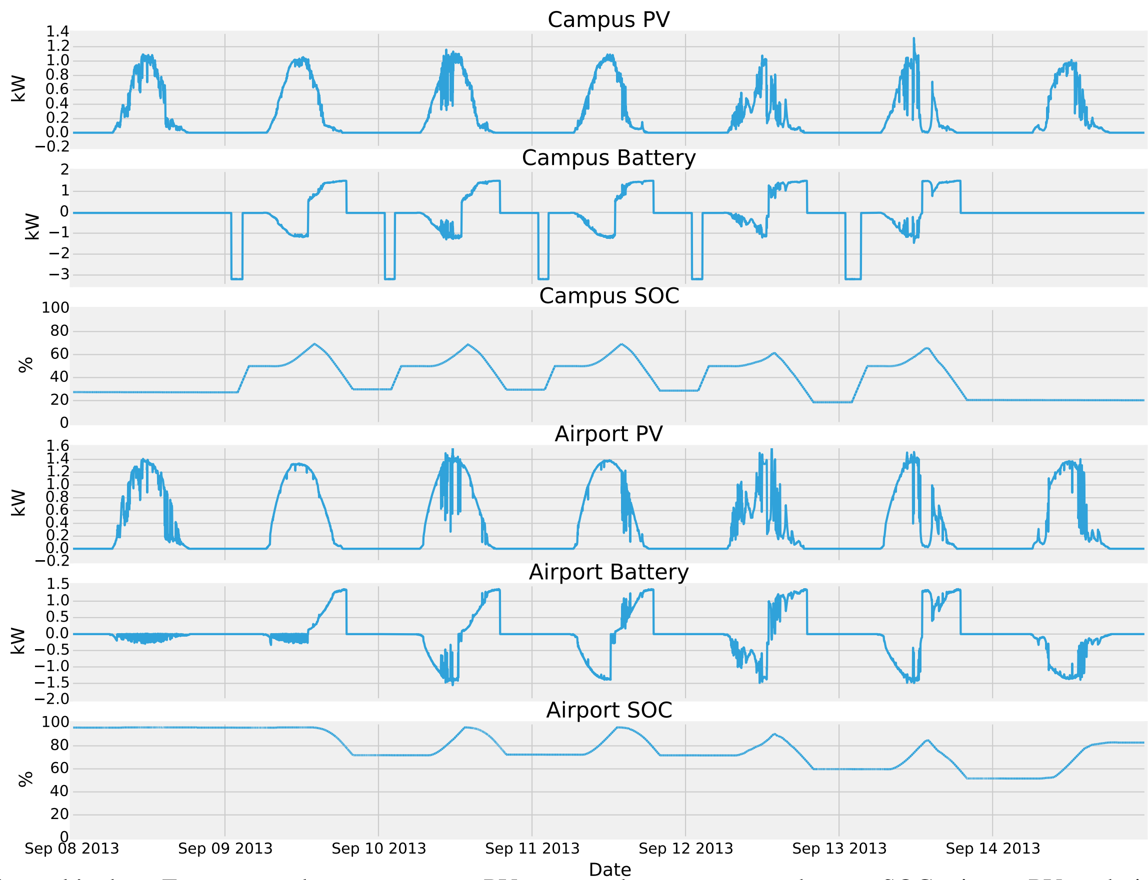


Fig. 3: A week's data. From top to bottom: campus PV, campus battery, campus battery SOC, airport PV, and airport battery, airport battery SOC. Note for batteries, reference power direction is assumed to be discharging.

large-scale data. This is not possible if Matlab is used. (ii) Through statistic analysis, PV daily energy versus environment variation is clearly shown. (iii) Battery capacity and efficiency degradation analysis is conducted using real-world data. The findings match degradation analysis in the literature.

II. Collected Data Format and Data Analysis Tools

One-minute interval data are collected. The measurements come from the four power meters installed at campus PV, campus battery, airport PV and airport battery. Approximately $525,600$ data points were collected for a whole year except data outages, which is shown in Fig. 2. Aside from ac power measurements, battery dc voltage, dc current and state of charge (SOC) are collected. The data are stored in spreadsheets as comma-separated values (csv)-based files.

TABLE I: Campus data outage records

Times	Start	End	Duration (min)
1	4/16/2016 19:34	4/16/2016 20:31	58
2	8/29/2015 20:20	8/29/2015 20:20	1
3	6/14/2014 1:10	6/14/2014 2:38	89
4	9/22/2014 9:14	9/22/2014 9:26	13
5	11/20/2014 13:07	11/20/2014 13:17	11
6	5/16/2013 12:30	5/16/2013 18:43	374
7	5/17/2013 22:10	5/17/2013 22:44	35
8	5/18/2013 3:43	5/18/2013 4:10	28
9	6/2/2013 0:11	6/2/2013 20:20	1210
10	10/23/2013 14:18	10/23/2013 16:42	145

Due to the large size of the data file, directly using Excel to make plots takes a large amount of time. In addition, automatic

TABLE II: Airport data outage records

Times	Start	End	Duration (min)
1	6/19/2016 14:01	6/19/2016 15:16	73
2	4/24/2015 15:06	4/24/2015 15:39	34
3	5/22/2015 10:13	5/22/2015 10:34	22
4	5/23/2015 14:31	5/23/2015 14:58	28
5	7/28/2015 10:04	7/28/2015 10:15	12
6	8/29/2015 20:20	8/29/2015 20:20	1
7	8/30/2014 5:45	8/30/2014 6:06	22
8	11/19/2014 21:06	11/19/2014 21:12	7
9	12/17/2014 16:13	12/17/2014 17:21	69
10	5/16/2013 6:55	5/16/2013 18:45	711
11	6/2/2013 0:13	6/2/2013 20:22	1210
12	8/12/2013 23:56	8/12/2013 23:59	4
13	8/13/2013 0:00	8/13/2013 0:26	27

plotting is difficult to be realized. In our data analysis work, we have conducted three tasks to make data analysis and plotting efficient.

- We have developed an SQL database to store four years' data in the database. Using query, we can then access the data fitting the query criteria. For example, we can list one week's data just by defining the time should be within a limit.
- Further, we have developed Python codes to access the database and make plots using Python module sqlite3 [4].
- Alternatively, we used Python module Pandas [5], [6] to directly access csv files and make plots using Matplotlib [7], [8].

The above tasks make data analysis efficient and possible.

III. Data Analysis Results

A. PV/Battery Operation

Fig. 3 presents the ac power data from September 8th (Saturday) to September 14th (Sunday) in 2013. Note the operation of campus battery and airport battery is to provide constant output power at 1:00 pm-7:00 pm. During each weekday morning, both batteries get charged using the PV power before 1:00 pm. Additionally, the campus battery gets charged in the early morning by electric power to ensure have enough energy for discharging operation in peak hours. Both those two batteries collaborate with PVs to provide constant power in the afternoon. There is no discharging scheduled for those two batteries on weekend.

Fig. 5 gives the campus site PV/Battery system outputs in summer and winter operation strategies. The total power (in red color) indicates that the combined system can effectively shift to provide constant power during peak hours in Summer (14:00-20:00) and Winter (06:00-10:00). The PV/Battery device would keep zero output if there was no need.

B. PV Daily Energy

Fig. 6 and Fig. 7 present the four-year PV daily energy for the campus PV and airport PV, respectively. The campus PV daily energy capture capability was improved after 2014. This is due to the removal of a tree at the site. Shades of the tree prevented the solar PV to absorb radiation.

The airport PV daily energy plot can be used to examine the weather impact on PV output. It can be clearly seen that in Tampa area, solar power is abundant in April and May. Storms happen in August and September days. Hurrican Irma formed on August 30 2017, and dissipated on September 13 2017.

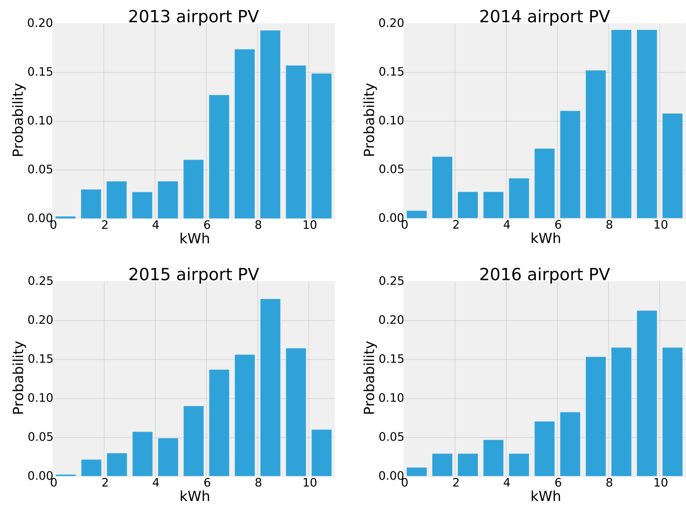
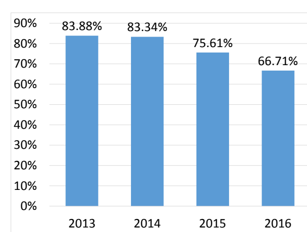


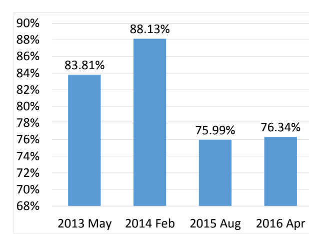
Fig. 4: 2013-2016 airport PV daily energy histograms.

The PV daily energy is computed from PV real-world power record. The record time interval is 1 minute. We approximately assumed the power was constant during each minute. Thus, we can sum up the power for a whole day to carry out the daily total PV energy through Python Pandas. The histograms in Fig. 4 can be easily plotted using Python's Matplotlib module.

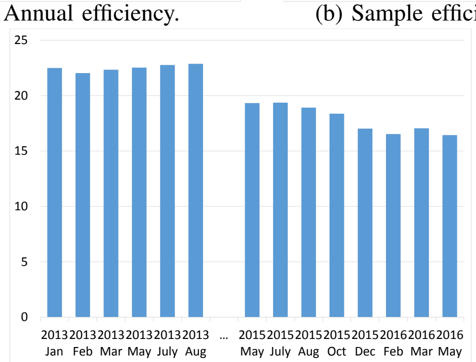
C. Battery Degradation Analysis



(a) Annual efficiency.



(b) Sample efficiency.



(c) Airport battery chargeable capacity in kWh.

Fig. 8: Airport battery degradation over time.

The battery degradation can be tested from two aspects. One is to check round-trip efficiency. Another is to check the

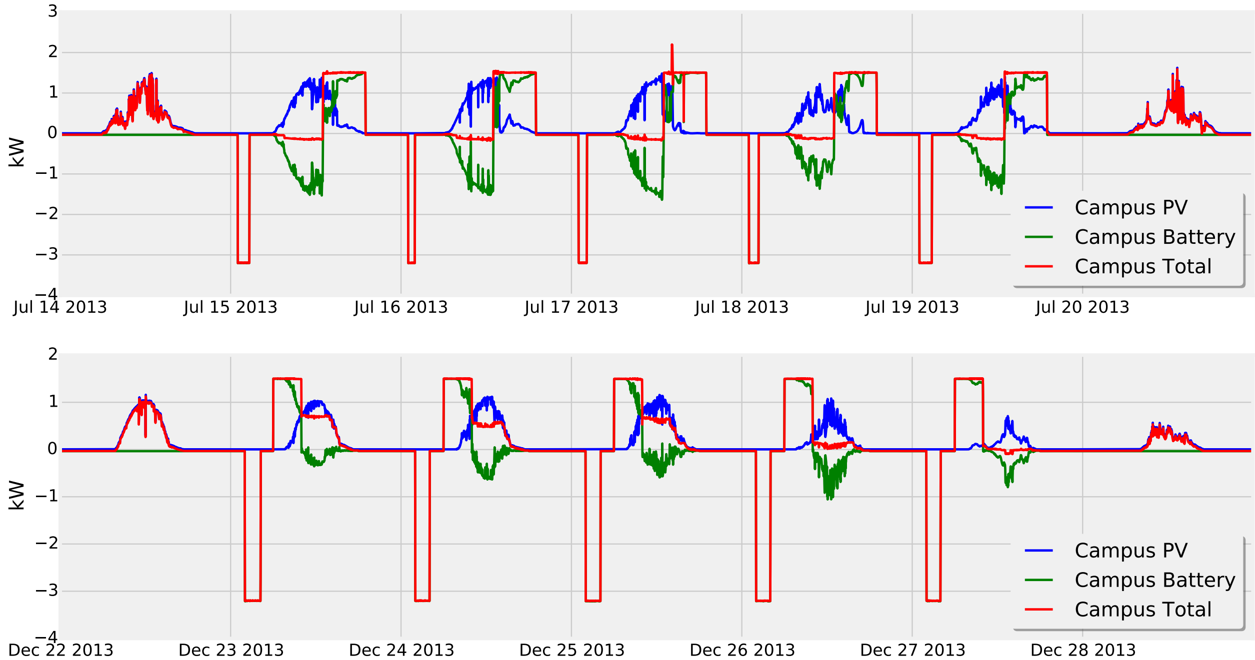


Fig. 5: Campus PV/Battery system summer (upper one) and winter (lower one) operations.

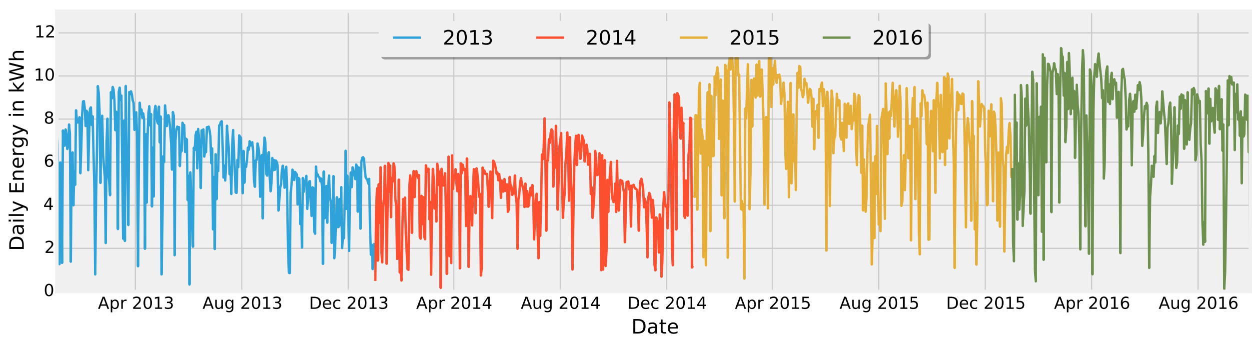


Fig. 6: 2013-2016 campus PV daily energy in kWh.

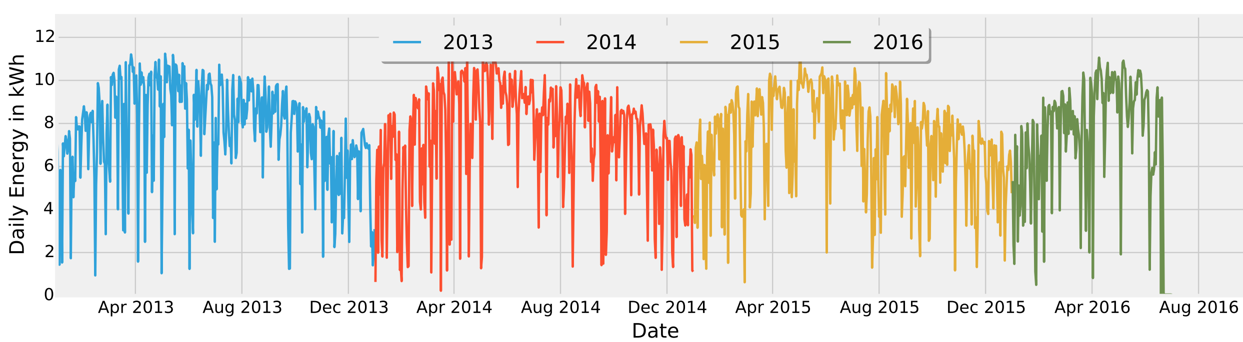


Fig. 7: 2013-2016 airport PV daily energy in kWh.

battery chargeable capacity over time. While battery capacity degradation due to aging is well known and experiments can be dated back in 2005 in [9] by MIT, few efficiency degradation experiments can be found, except a recent publication on small lithium-iron cell at 2.3 Ah [10]. In this paper, both the degradation analysis on efficiency and capacity for a 20 kWh battery will be presented.

We use annual efficiency and sample efficiency to check battery round-trip efficiency. First, each year's annual efficiency is calculated through the battery output power spanning a whole year. We can treat one year as a long-term round-trip since the beginning SOC is closed to ending SOC for each year.

The percentage of data outage is less than 1% so that we can ignore them. The ratio of the whole year's discharged energy to charged energy is the annual efficiency, shown in Fig. 8a. Overall, we see a decrease in round-trip efficiency.

On the other hand, one fully charging/discharging cycle sample is extracted from each year to test sample efficiency. The data is listed in following TABLE III. Here, SOC should start from very small value and rise to nearly 100%, then drop back to a similarly small number. The sample period in 2013 is detailed in Fig. 9. Fig. 8b represents the efficiencies computed from 4 samples in TABLE III.

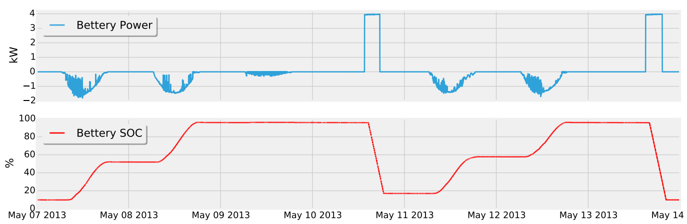


Fig. 9: Airport battery sample period in 2013.

TABLE III: Airport battery round-trip samples

Year	Sample Period	SOC Range (%)	Efficiency
2013	May.7—May.14	9.8—96.0—9.8	83.81%
2014	Feb.15—Feb.28	15.9—99.3—14.8	88.13%
2015	Aug.1—Aug.8	1.5—99.4—1.5	75.99%
2016	Apr.17—Apr.22	1.4—99.3—1.6	76.34%

Furthermore, we extracted data for cycles with small SOC deviations (e.g., 10% to 20%) for each year to investigate efficiency degradation versus SOC level. Take Fig. 10 as an example, we use a cycle 68% $\rightarrow$ 90% $\rightarrow$ 68% to compute the efficiency at SOC level at 79%. The 4-year Airport battery efficiency at different SOC levels are presented in Fig. 11. We can observe the efficiency is almost constant when SOC level is less than 60%. The efficiency degrades with SOC increasing in the high SOC region. In addition, the efficiency degrades with battery aging. Efficiencies in 2015 and 2016 are lower than those in 2013 and 2014. That can also explain the annual efficiency degradation and full depth cycle efficiency degradation in Fig. 8.

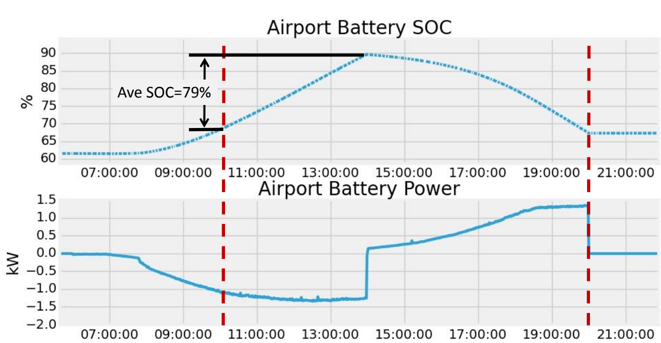


Fig. 10: Airport battery SOC cycle sample with small SOC deviation.

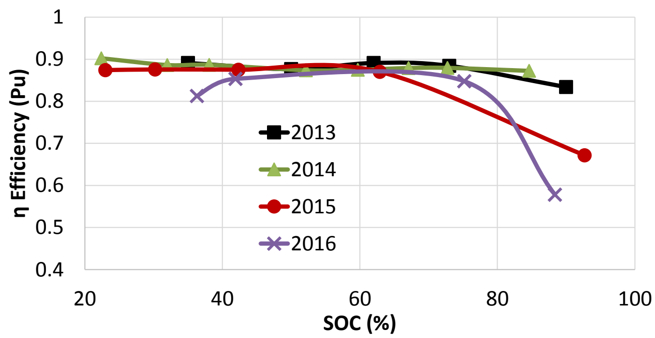


Fig. 11: Airport battery efficiency at different SOC levels.

The variation of efficiency due to SOC level and battery aging can be explained by the equivalent circuit model of Lithium Iron battery proposed by Liaw in [9] and applied by other researchers, e.g., [11]. It has been recognized that the equivalent resistance of a battery is related to SOC and their relationship is nonlinear [11]. The overall cell resistance increases when a cell is aging and the SOC increases. This explains efficiency degradation with SOC increasing in high SOC region and efficiency degradation with battery aging.

IV. Conclusion

This paper presents data analytics based on real-world 1.6 kW PV/20 kWh Battery systems. Besides statistic analysis, e.g., daily PV energy over four years, the airport battery degradation analysis has been conducted through round-trip efficiency computing and total chargeable capacity computing.

V. Acknowledgement

We would like to acknowledge George T. Gurlaskie of Duke Energy for his support.

References

[1] M. Zhang, Z. Miao, and L. Fan, "Battery identification based on real-world data," in *Power Symposium (NAPS), 2017 North American*. IEEE, 2017, pp. 1–6.

[2] Y. Xu, Z. Miao, and L. Fan, "Deriving arx models for synchronous generators," in *North American Power Symposium (NAPS), 2016*. IEEE, 2016, pp. 1–6.

[3] B. Mogharbel, L. Fan, and Z. Miao, "Least squares estimation-based synchronous generator parameter estimation using pmu data," in *Power & Energy Society General Meeting, 2015 IEEE*. IEEE, 2015, pp. 1–5.

[4] M. Owens and G. Allen, *SQLite*. Springer, 2010.

[5] W. McKinney, "pandas: a foundational python library for data analysis and statistics," *Python for High Performance and Scientific Computing*, pp. 1–9, 2011.

[6] W. McKinney *et al.*, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, vol. 445. Austin, TX, 2010, pp. 51–56.

[7] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in science & engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[8] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython.* " O'Reilly Media, Inc.", 2012.

[9] B. Y. Liaw, R. G. Jungst, G. Nagasubramanian, H. L. Case, and D. H. Doughty, "Modeling capacity fade in lithium-ion cells," *Journal of power sources*, vol. 140, no. 1, pp. 157–161, 2005.

[10] E. Redondo-Iglesias, P. Venet, and S. Pelissier, "Efficiency degradation model of lithium-ion batteries for electric vehicles," *IEEE Transactions on Industry Applications*, 2018.

[11] A. Millner, "Modeling lithium ion battery degradation in electric vehicles," in *Innovative Technologies for an Efficient and Reliable Electricity Supply (CITRES), 2010 IEEE Conference on*. IEEE, 2010, pp. 349–356.