Detecting & Forecasting Solar DC Field Losses

Submitted by:
Taylor Ballard
Anthony Funk
Jimmy Nyanwapolu
Dhara Patel

Project Report

EMGT 5963: Energy Systems Planning
Fall 2019
Systems Engineering & Engineering Management Department
The University of North Carolina at Charlotte

Table of Contents

Executive summary

The decarbonization revolution is beginning for many utilities in the United States. As this occurs, it is inevitable for our dependency on coal and natural gas to reduce. Solar and wind production has become more attractive to utilities due to dropping costs and growth has increased exponentially over the last decade. Due to the nature of renewable energy production based on resources that are not guaranteed, efficiency of generation plants is much lower than traditional forms. As renewable generation is not always available, it is critical for plants to operate at maximum efficiency. Data analytics can help assess the performance efficiency of solar assets and identify underperformance of the solar panels.

Commercial size solar sites experience many types of lost energy production. The most prominent loss of generation comes from the DC solar field itself. On sites that span hundreds of acres, common DC field issues include bad panels, connections, ground faults, failed combiner box fuses, and failed inline fuses. Managing failure modes throughout sites of this size can be challenging. Data driven techniques must be used in order to effectively reduce availability lost energy (ALE) by charactering failures and optimizing work performed. This project will aim to reduce solar site ALE by applying new techniques to identify and prioritize field failures. We will also apply techniques learned in class to forecast solar site output based on multiple variables. We can use Forecasting Methods like Multiple Linear Regression (MLR), Artificial Neural Network (ANN) and Support Vector Machine (SVM) to forecast the String Current for Inverters.

Introduction

The cost of silicon solar panels has decreased significantly year over year, while efficiency has improved dramatically. Technology improvements with attractive costs have created a sharp increase in development for utilities. In addition, many countries have offered incentives and tax credits for renewable energy projects. Over the last decade, the cost of solar has dropped by over 70% making solar energy generation competitive with traditional forms of energy production. Dropping costs have driven expansion, resulting in an annual growth rate of 50% since 2006. Expansion of emerging technologies brings about new challenges, such as managing site efficiencies. Data analytics has been used in most major industries such as finance, healthcare, aerospace etc. The application of data analytics can help optimize the solar power generation and forecasting.

NextEra Energy is one of the largest renewable energy companies in the world. Recently, the company has significantly expanded its generation from wind and solar. Florida Power & Light (FPL), a subsidiary of NextEra Energy, recently committed to the state of Florida the 30-by-30 initiative. The initiative calls for 30 million solar panels by 2030, propelling the state of Florida to be one of the largest producers of Solar Energy in the world. FPL has 17 active solar sites that are capable of outputting 75MWs each. This number is set to double in the next two years alone. The project will use data from these generating facilities to look for a solution to detect DC field losses and forecast future generation.

Problem Statement

Large scale commercial solar sites have no method to track, analyze and account for losses that occur in the solar DC field. Solar sites already have low capacity factors due to sunlight during only day-time hours. Losses that occur in the DC field significantly impact the generation potential during these hours. Companies are forced to spend large sums of money for contractors to walk down hundreds of acres of land to troubleshoot, locate, and repair DC field issues. Through data analysis, we hope to identify solar panel circuits that have failed and are not providing maximum output. The burden of handling the end-of-life losses of PV system, and recycling of massive degraded panels in the next 20 to 40 years of PV energy generation demands the need for long term system loss and load forecasting

Literature Review

To convert sunshine to electricity, photovoltaic (PV) systems are subject to many sources of energy loss (losses). Understanding the mechanisms of DC sites losses and how to minimize them is very essential to optimizing and accurately predicting the performance of PV system. To determine the magnitude of PV losses, there are various elements that must be considered, and they are the module and balance of system (BOS) technologies, location of the installation, array and inverter configurations, and other factors. Some solar plant installations accept higher losses to save on the initial capital investment or operations and maintenance (O&M) costs. Losses also vary based on time of day, year, and location. An example is, during high ambient temperatures, PV modules may generate 5-15% lower output than at standard test conditions.

Description of PV system losses are based on power conversion, energy flow from the full energy of the sun to the delivered energy at the utility's point of common coupling. Additionally, losses also stem from damaged or malfunctioning equipment, such as broken connections, tripping events and maintenance issues. According to a December 2013 Assessment of Photovoltaic System Losses report, by Electric Power Research Institute (EPRI), the major categories of losses identified are classified into:

- Solar Resource and the Environment: solar radiation, sun path & angle of incidence, etc
- Photovoltaic Modules: the material properties, conversion efficiencies
- Arrays: Orientation & Tilt (adjusting the angle of incidence)
- Inverters: Connection topologies, threshold & standby losses
- Interconnection: point of common coupling (utilities, Residential or commercial)
- Operation and Maintenance: Corrective, conditional, and preventive

According to a 2011 Desert Research Institute report by Daniel Shapiro, C. Robbins, and P. Ross, natural degradation of solar cells, solar panel end of life, and recycling options are considered loss factors associated with DC site and its corresponding O&M cost. Solar cells degrade naturally overtime, irrespective of the environment they are in, and it is a normal experience for every solar cell once in operation. Silicon (Si) panels including both mono crystalline and poly crystalline degrades on average rates of 0.7% a year, with a median value of 0.5% a year. Manufacturing companies such as Jinko Solar, Solar World, and Q Cells offer warranties if degradation rates exceed certain amounts. For Jinko Solar, Solar World, and Q Cells, user qualifies for panel replacement or compensation if degradation values exceed 0.8%, 0.7%, or 0.6% respectively for each company. This shows that a higher quality panel will have less natural degradation (2011, Desert Research Institute data)

Calculation of accurate returns on solar plant investment therefore requires accounting for the elements of DC field losses through developed procedures. This would ensure plant efficiency and cost effectiveness in both the short and long run. Solar panel end of life and its recycling option after plant is decommissioned are long term loss effects and O&M issues of DC sites and are therefore not included in the field loss procedures. For residential systems this analysis method may not be as much of a problem as it is with commercial due to not being in the same range of 1 MW systems that consist of approximately 4000 panels.

The project team anticipates that analysis of the site DC data collected from NextEra Energy's PV system will indicate how some if not all of the system loss categories described herein contribute to the predictive model developed, with the aim to enabling PV system optimize performance through accounting for loss forecast. Building a model to forecast and identify the losses within the systems will allow for a swifter and more in-tune maintenance regimen for the PV sites creating a cost effective and targeting operation and maintenance procedure for the company.

Data Collection, Techniques, and Methods

I. Detecting DC field Losses

To have a better understanding of the data, a brief description of site geography is needed. Each site has 80 inverters (~1MW each) that are split into 4 blocks for naming convenience. The blocks are 11, 12, 21 and 22. The first two numbers of the inverter's name is the block in which it is located. Inverters are then grouped into skids, which is a concrete palate with two inverters. These are denoted the alpha and bravo inverter. A block would be named 1101A, 1101B, 1102A so on and so forth.

Figure 1 An Inverter Skid

Figure 2 Layout of site from which we collected data, skids are represented by blue squares

Data will be collected from solar sites in the west region of Florida. The initial sample size was just one site, but the sample size can be expanded to 7 sites if more data is needed. Variable metalogical data was collected such as irradiance, humidity, reference cell temperature, wind speed, and air temperature from site meteorological stations. In addition, all solar string currents (8 strings per inverter) was be collected for all 80 inverters on site. The data was pulled in one-minute increments for all of October 2019. In order to more accurately forecast, we also pulled all met station data for November 2019 to validate our model.

The string data is recorded from current transformers (CTs) that are placed on each string at the bottom of each inverter DC cabinet.

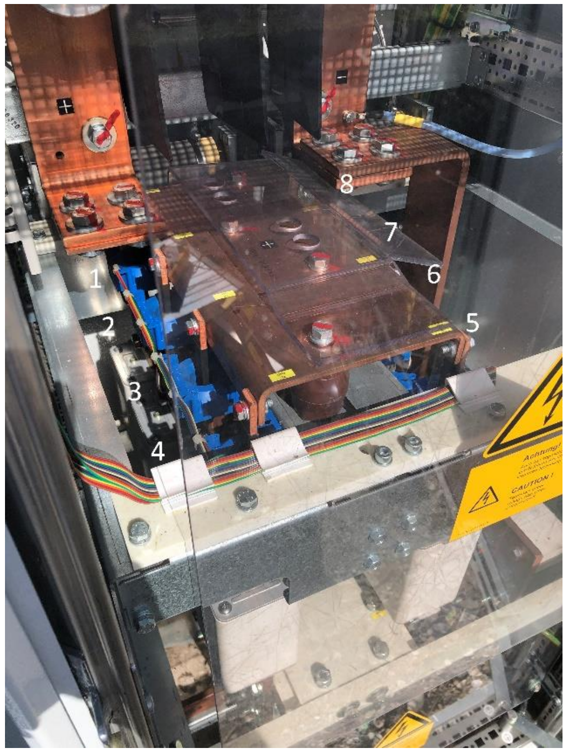

Figure 3 8 Blue String Current CTs

After collecting all the data, we looked to find the met station variable with the most direct correlation to string current output. This variable was found to be irradiance. We modeled the relationship between the predictor (irradiance) verses the continuous response (output string current) and found the two variables have an R squared value of 95.9% with a standard error of regression of 9.45. We also took note of instances in the data where the inverter output was 0 while the irradiance was >0, indicating the inverter was in forced outage. These data points likely pull down the still-strong R squared value.

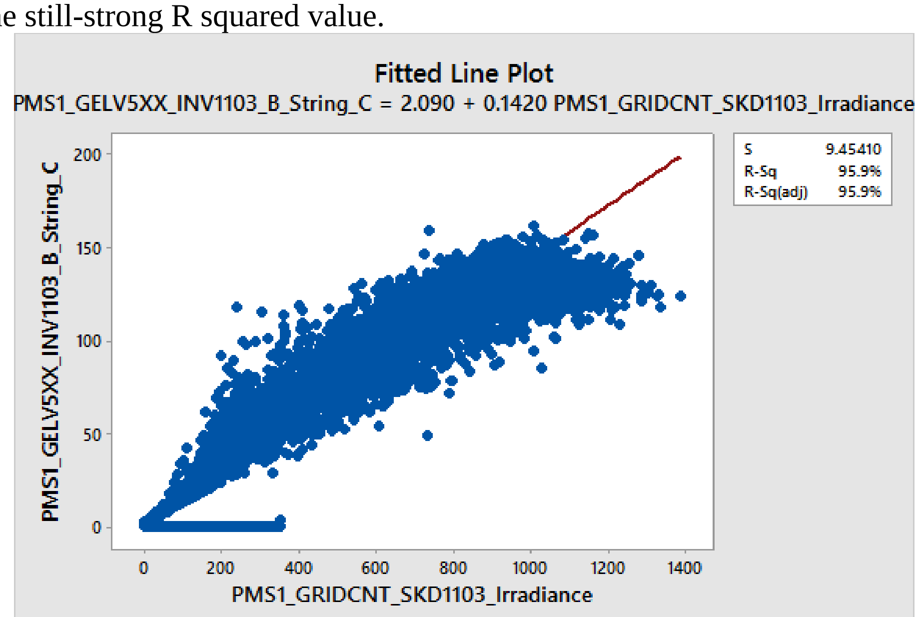

Figure 4 Regression model for irradiance and string current

Next, we aimed to find a day where sun intensity was relatively consistent throughout the day. We wanted our string current data to be as clean as possible while aiming to look for deviations in current output. The team used Microsoft Excel to create a formula that calculates when irradiance values fall lower than a previous value (from 0700 – 1245 hours) and calculates when irradiance value falls above a previous value (from 1245 – 1800 hours). Similar to the methods taught in class, we extracted date and time information to be used in the calculation. In theory, a perfect blue-sky day would have an ever-increasing irradiance value until mid-day, followed by a steadily decreasing irradiance value until sundown. The excel calculation looks for deviations. The curve would look like a perfect parabola with no clouds in the sky throughout the day. Moderate cloud cover would cause quick shifts in irradiance, and these are the days we wish to avoid. Using this formula, we were able to exclude days with high variance in irradiance and locate a near perfect sunshine day on 11/30/2019 to run our data analysis.

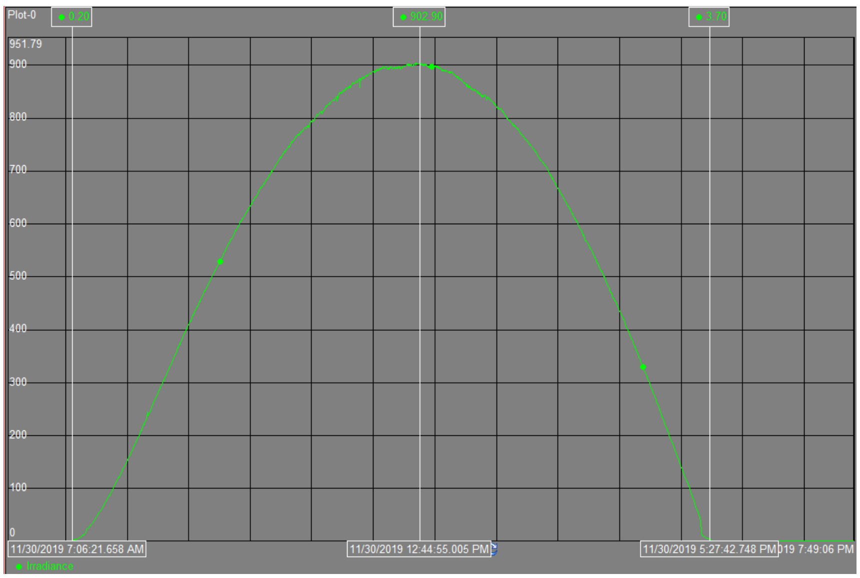

Figure 5 Near perfect clear sky day to continue our string current analysis

As computer power is limited and our data is millions and millions of data points, we took a sample size of five inverters and their associated 8 string currents to locate any deviation in mean from one string to another. On inverter 1104B, we found exactly what we were looking for. The string current means were calculated on the table below.

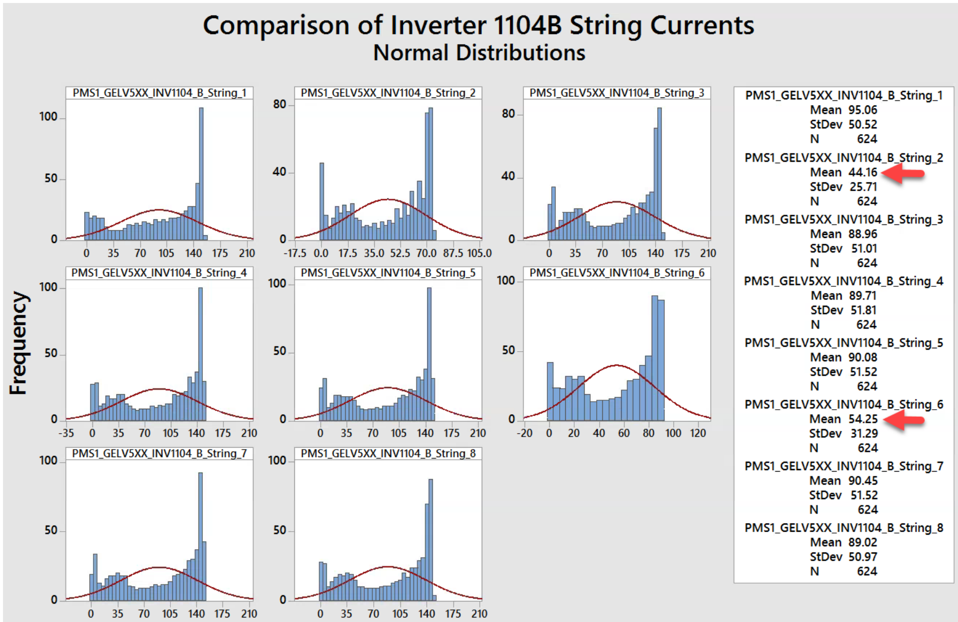

Figure 6 1104B String Current Data Distributions, Means, and Standard Deviation

From the above statistical analysis, one can deduce that strings 2 and 6 have failed circuits that are dragging the mean output down in comparison to the other strings. Next, we plotted the data over time.

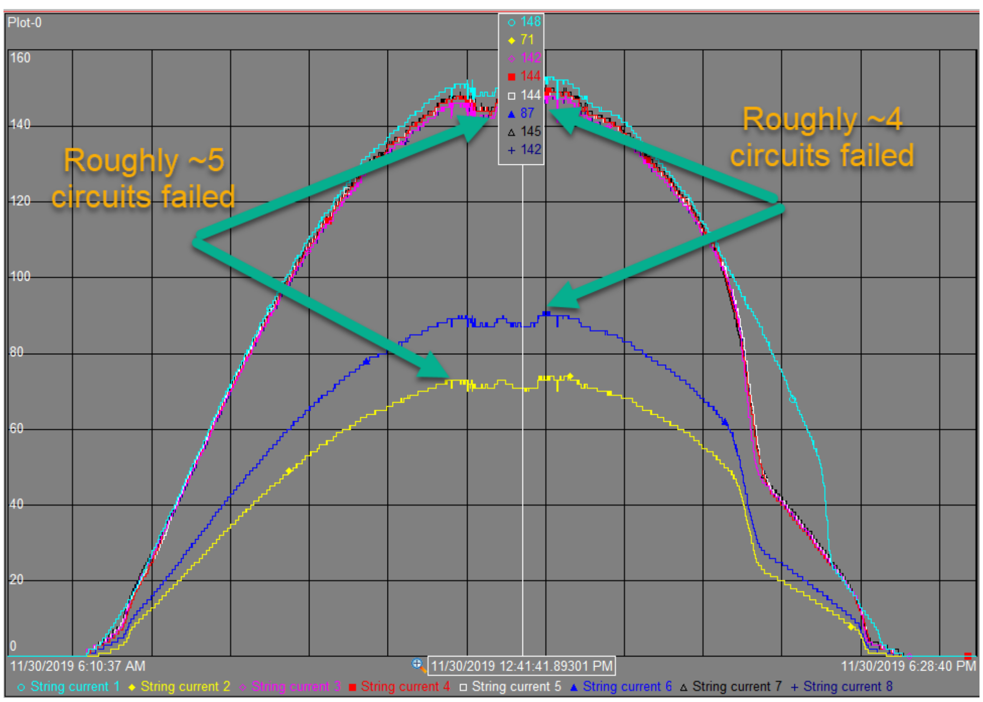

Figure 7 String current outputs plotted over daylight hours on 11/30/2019

Each circuit generates roughly 15A at maximum power. Taking the maximum power point from the day, we can deduce that string 2 has roughly 5 failed solar circuits and string 6 has roughly 4 failed circuits. Assuming data can be fed into a complex algorithm, it should be possible to not only find DC field losses, but associate a severity to those losses! Results in the field are discussed in the next section.

II. Forecasting Solar Site Generation

There are different Forecasting Techniques like Multiple Linear Regression, Artificial Neural Network, Support Vector Machine, XGBoost, ARIMA, Random Forest. For this Project MLR and ANN is used to forecast String Current. DC Field losses has been found using different string currents of the same inverter. Forecasting was done for String Current 1 to 8 for INV1101A, 1103A and 1104B.

Multiple Linear Regression:

Multiple linear regression is the most commonly used modeling technique.

MLR is used to model the linear relationship of string current and other factors such as air temperature, reference cell temperature, humidity, wind speed and solar irradiance.

For the MLR data of the October Month is divided in the 80-20 ratio for training and testing the model. 80% of the data is used to train the model. And 20% of the data to test the model.
We also tried with 50-50, 70-30 and 75-25 ratio of training testing and second order temperature values. After that used the best variables and 80-20% ratio to predict the value of string current for month of November.
Regression equation (General):
String Current = a + b (Air Temperature) +c (Reference cell Temperature) +d (Humidity) +e (wind speed) + f (solar irradiance)

Artificial Neural Network:

The ANN is used to perform nonlinear modeling for the relationship between the string current and weather variables.

The output of an artificial neural network is some linear or nonlinear mathematical function of its input. The inputs may be the outputs of the other network element as well as actual network inputs. Multilinear Perceptron is used to predict the string current_1 to 8 for INV1101A, INV1103A and INV1104B.

Results and Technical Discussion

I. Detecting DC field Losses

With the discovery of several failed circuits on strings 2 and 6 of Inverter 1104B, results were verified accurate in the field with the following steps.

1. Determine that an issue in the DC field exists. The team successfully calculated the strings do not have similar outputs during the same time of day.
2. Using data algorithms, determine how significant the loss is and whether it is cost effective to walkdown, troubleshoot and fix. The team was able to roughly estimate how many circuits have failed. An additional analysis will need to be completed in a future project to determine what amount of loss is cost effective to fix.
3. Field troubleshooting was initiated starting by determining if the issue was affecting all circuits of a string combiner box or only isolated circuits. If the entire combiner box is down, that string will read 0A, a significant loss. Each combiner box feeds into the string

CT shown in the picture above. We know based on our data none of the strings tied to this inverter read 0A. The combiner box pictured below is string 6 of inverter 1104B.

Figure 8 Combiner box where solar panel circuits meet to form a string

4. Next, each solar circuit that feeds the combiner box will need to be evaluated. The first step is to check the 30 amp fuse (the horizontal black fuse holders) shown in the picture below. Each fuse carries the load of a single solar circuit that is combined in the box to form a string.

Figure 9 Inside a combiner box, showing solar circuits coming in tied to a 30A fuse

Circuits are named P01 through P10 in the combiner box. Immediately discovered, P05 had a burned-out fuse holder. The fuse on P01 had also blown out, identifying 2 of the 4 suspected bad circuits immediately.

Figure 10 30A fuse holders showing where solar circuits combine into a string, P05 burned out

5. All the remaining 30A fuses tested successfully. After troubleshooting the 30A fuses, the next step will be to check the 15A inline fuses that are throughout the solar field. Each circuit will need to have its current checked, looking for a circuit that features either half or zero current compared to other solar circuits. If a circuit is discovered in

this condition, it is likely the 15A inline fuses are the culprit. Discovered on P07, both inline fuses were found to be blown.

Figure 11 Inline fuses located directly under solar panels, both defective

6. Finally, to identify the last failed circuit the entire string was walked down. On P09, the worst failure mode possible was discovered - failed solar panel connections and wiring.

Figure 12 Failed solar panel connection

Figure 13 Failed wiring

Our data analysis was proven accurate, finding 4 failed circuits on string 6 of Inverter 1104B.

II. Forecasting Solar Site Generation

As Discussed Earlier We have used MLR and ANN(MLP) to forecast the string current. Accuracy of the model is calculated using Root Mean Square Error.

RMSE for all string current is as below.

MLR:

Table 1 RSME for MLR Analysis

	INV1101A (%)	INV1103A (%)	INV1104B (%)
StringCurrent1	11.62	8.67	10.46
StringCurrent2	11.79	8.26	4.80
StringCurrent3	11.82	8.65	9.70
StringCurrent4	11.98	7.98	9.55
StringCurrent5	12.20	8.74	9.27
StringCurrent6	12.20	8.24	6.82
StringCurrent7	12.37	8.10	8.92
StringCurrent8	11.36	9.79	8.63

ANN

Table 2 RSME for ANN Analysis

	INV1101A (%)	INV1103A (%)	INV1104B (%)
StringCurrent1	9.36	6.49	7.44
StringCurrent2	9. 82	6.94	3.10
StringCurrent3	9. 80	6.73	7.71
StringCurrent4	9.67	6.00	7.83
StringCurrent5	9.82	6.47	7.80
StringCurrent6	10.00	6.84	5.16
StringCurrent7	10.02	6.66	6.84
StringCurrent8	9.67	7.80	6.28

For INV1104B String Current 2 and String Current 6 have less current than all other inverters. It seems there are DC losses with these 2 strings. Result has been verified with the original data.

Summary and Conclusion

O&M costs for solar sites is very high if companies intend to keep the DC field performance close to ideal condition. For solar power generation to be a major player and compete with fossil generation costs, solar O&M costs must be reduced. From our analysis of the data, it appears feasible to not only detect DC field losses but predict their significance. The company can use highly advanced algorithms to perform this analysis in real time and alert solar site operations when field failures occur.

In addition, solar site operations teams can create a forecast based on the historical data of weather and string currents and NWP data of the weather for future predictions of string currents to predict DC losses. In addition, we can use historical data of Faults to predict the other Faults in solar site and we can replace the component or repair it during the time of less power production hours. Which helps to saves lots of dollar amount to the company.

References

Dunlop, James. Photovoltaic Systems. Orland Park: American Tech Publishers, 2012.

Fobes, Wes. 2013. "Assessment Of Photovoltaic System Losses". Technical Brief - PV, Reliability, Design, O&M, Solar Program. Palo Alto, California: Electric Power Research Institute.

Shapiro, Daniel. 2019. "Solar PV". DRI.Edu. https://www.dri.edu/images/stories/editors/receditor/Solar_PV_Article.pdf, retrieved on November 16, 2019.

Twidell, John W., and Tony D. Weir. Renewable Energy Resources. 2nd ed. New York: Taylor & Francis, 2006.

IEEE Std 1547.3 TM Guide for Monitoring, Information Exchange, and Control of Distributed Resources Interconnected with Electric Power Systems 26 Uriarte, Fabian M., Robert E. Hebner. Impact of Residential Photovoltaic Generation and Electric Vehicles on Distribution Transformers: Progress Report# 1."April 08, 2013 (pages vii, 50) http://www.pecanstreet.org/wordpress/wp-content/uploads/2013/04/PSRI_CEM_white_paper-Impact_of_Residential_PV-UriarteHebner-2013-04-08.pdf. 27

Jansson, P.M., U.K.W. Schwabe, A. Hak. "Large-Scale Photovoltaic System Design: Learning Sustainability through Engineering Clinics." (Tables 3, 4) http://www.rowan.edu/colleges/engineering/clinics/cleanenergy/pv/papers/pdf/files/paper5.pdf.