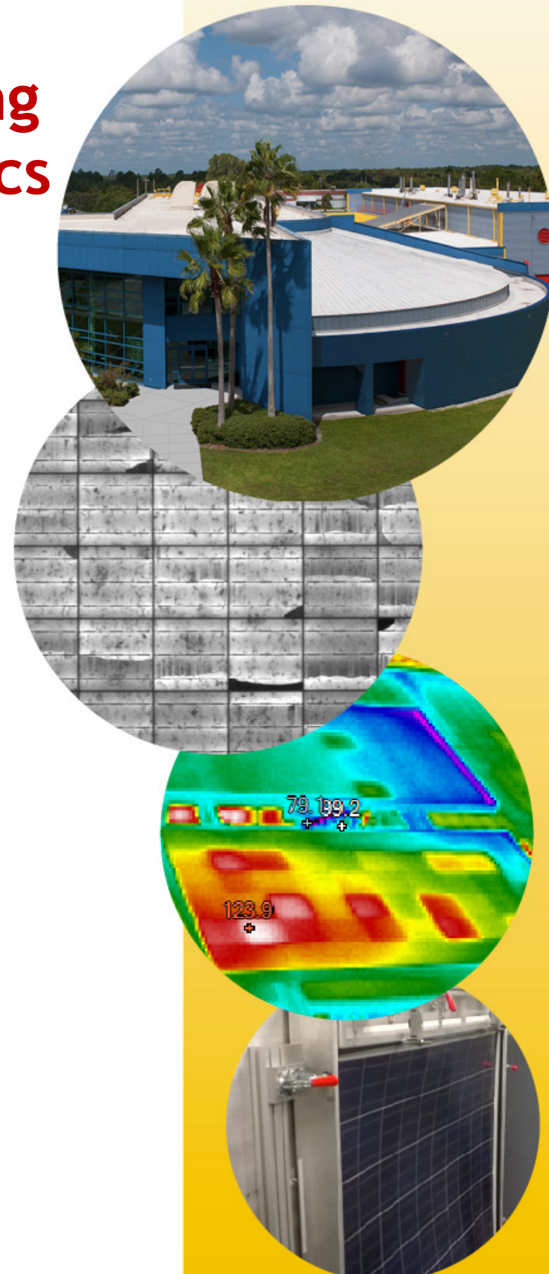
UCF | FSEC Energy Research Center

UNIVERSITY OF CENTRAL FLORIDA

Best Practices Guide for Utility-Scale PV Monitoring with Intelligent Diagnostics Using String-Level I-V Curves and Machine Learning



Report for public release

https://publications.energyresearch.ucf.edu/

July 26, 2022

Submitted to

Department of Energy
1000 Independence Ave., SW
Washington, DC 20585
Award No. DE-EE0008157

Authors

Manjunath Matam, manjunath.matam@ucf.edu
Hubert SEIGNEUR, *seigneur@ucf.edu*
Ryan Smith, ryan.smith@pordis.com

1679 Clearlake Road
Cocoa, FL 32922-5703
Tel: 321-638-1000 • info@fsec.ucf.edu

Best Practices Guide for Utility-Scale PV Monitoring with Intelligent Diagnostics Using String-Level I-V Curves and Machine Learning

Today, 121 GW Photovoltaic (PV) plants are installed and operated in the USA, which is 11% of the total installed power capacity. Of these, 75% were installed in the last five years, the highest growth rate among all power sources. Therefore, PV plants need to be monitored to confirm that plant power output, yield, and life expectancy are within the expected limits, which ensures the PV owner's financial goals are attainable. Industry-standard monitoring methods collect and analyze the inverter time series data, typically currents and voltages. However, inverter level monitoring lacks granularity, detection sensitivity, and localization. Therefore, plant operators must resort to occasional aerial infrared (IR) imaging to detect malfunctioning strings and modules. However, IR imaging cannot provide failure root cause, degradation rates, nor detect mismatch losses, which can grow to be as large as 20% over time [1]. A potential solution is performing advanced diagnostics at the string or module level through in-situ I-V characteristic measurements, from which a lot of information can be extracted [2]. Anticipated advantages include continuous automated failure diagnostics performed on strings and modules, improved localization of faults, more effective preventive maintenance to reduce downtime and increase energy production, in-situ degradation rate estimation for product warranty monitoring and reduction of LCOE and risk due to increased availability and power generation as a result of proactive O&M activities.

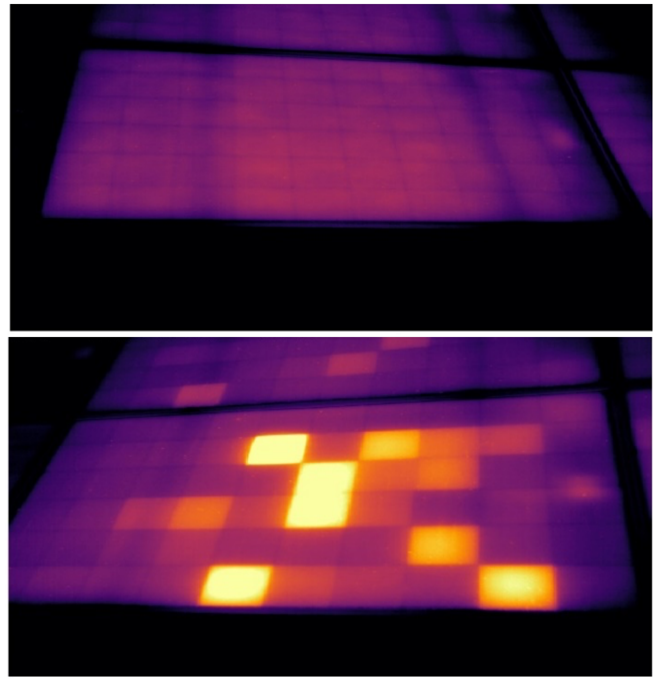


Figure 1: The popular infrared imaging technique is good at detecting hotspots but under certain conditions only. The top and bottom are IR images of a string showing a specific PV module. The module showed no signs of hot spots when string was operated at MPPT. But, the same module showed signs of checkered pattern hotspots when operated at off-MPP. *Photo was taken at FSEC test site.*

About the project

The project titled “LCOE reduction through proactively optimized monitoring of PV systems” was sponsored by the US Dept. of Energy through the award DE-EE0008157. The objective is to demonstrate the value proposition for a high-resolution monitoring system (HRMS) with diagnostic-prognostic capability and determine its impact on LCOE. More details about the project, publications, reports, data, and proposed fault detection algorithm can be accessed at https://publications.energyresearch.ucf.edu/

This guide is intended to help PV plant owners, engineering, procurement, and construction (EPC) firms, monitoring service providers, and other industry stakeholders by providing important findings and lessons

[1] W.G. Shin et al, Currentflow analysis of pv arrays under voltage mismatchconditions and an inverter failure ApplSci, 9 (2019),10.3390/app9235163
[2] J. Walters, et al"Experimental Methods to Replicate Power Loss of PV Modules in the Field for the Purpose of Fault Detection Algorithm Development," IEEE 46th PVSC, Chicago, IL, USA, 2019, pp. 1410-1413

learned in designing a monitoring system to detect PV faults and is the culmination of a project in which the simulation, hardware, and indoor and field experiments were conducted at a test site and a utility PV plant. In addition, this guide offers research findings, information, and recommendations for the earliest detection of PV faults, advanced monitoring practices, cost-economic benefits, LCOE analysis, and comparing different monitoring approaches and their impact on fault detection. Also, the checklists mentioned throughout the guide help the PV monitoring stakeholders with critical information when designing and installing string-level I-V curve-based monitoring equipment.

Levels of Monitoring

Four levels of monitoring can be implemented at a PV plant. Of these, measuring DC and AC parameters at the PV array (or inverter) level is a conventional approach, and is required by the IEC-61724 standard. The remaining three are options available to PV plant stakeholders and are not well documented, researched, and are yet to be fully commercialized. The details of the four monitoring levels are:

1. Inverter DC and AC parameters (time series): Apart from performing DC-to-AC energy conversion, inverters are fully equipped to measure and report the PV array's input DC and output AC parameters to the PV plant's operators. It is time series data. The measurement frequency is a choice of the PV plant. For instance, IEC-61724 categorizes PV plants into three categories: A, B, and C, each with a measurement frequency of <1 min, 1-15 min, and 15-60 min, respectively. Irrespective of the measurement frequency, inverter data provides little or no information about the faults in the PV array. It is the lowest-cost monitoring configuration as the data acquisition system and measurement devices are typically connected only to the inverter. It is the most widely accepted monitoring setup. Using such a system, PV plant operators raise a maintenance ticket only when they notice a minimum 10% loss in the DC or AC parameters. Since 2010, commercial inverter manufacturers have been introducing their versions of "smart inverters," with varied differences between manufacturer and model. Today, equipping "smart inverters" with state-of-the-art capabilities to measure the characteristic I-V curves of the PV array or sub-array is expected. However, with a limited number of I-V curve tracing inputs per inverter, it may not be able to measure the I-V characteristics of individual strings and still lacks some granularity and localization.
2. Combiner current transducer (time series): A combiner is an electrical junction of the strings, usually 12 to 25, feeding current to a recombiner or inverter. A current measuring device is integrated into the combiner with the intent to measure and collect combiner current and analyze the information to detect faults in individual strings or a group of strings. It is a low-cost monitoring setup. The measurement frequency differs based on the category of PV plant. Plant operators raise maintenance tickets when they notice a minimum of 8-10% loss in the current.
3. String voltage and current (time series or I-V curve): Monitoring each string is a state-of-the-art approach to effectively detect string level faults. Although a usual practice is monitoring the MPP current and voltage parameters, this more advanced approach involves measuring the characteristic I-V curve of the string. The former needs no special equipment except the integration of the devices. However, the latter approach needs additional arrangements, that is, a 1-2 second isolation from the PV array with the help of switches for momentary disconnection and reconnection. The I-V curve measurement is usually quick and fast, taking less than a second. However, compared to the previous methods, it requires measuring hardware installed at the string/combiner level. Still, it offers the advantage of near real-time detection of faults in the string leading to early prevention or correction measures and low losses. **This project's extensive simulation, experimental tests, cost-benefit, and LCOE analysis showed that string-level monitoring is optimal for detecting and locating the most PV faults without requiring hardware on every single module.**
4. Module voltage and current (time series or I-V

curve): Measuring module level current and voltage or characteristic I-V curve is the next level in the monitoring. It is expensive but maybe the default approach for micro-inverter applications as they are inherently present in module-level devices.

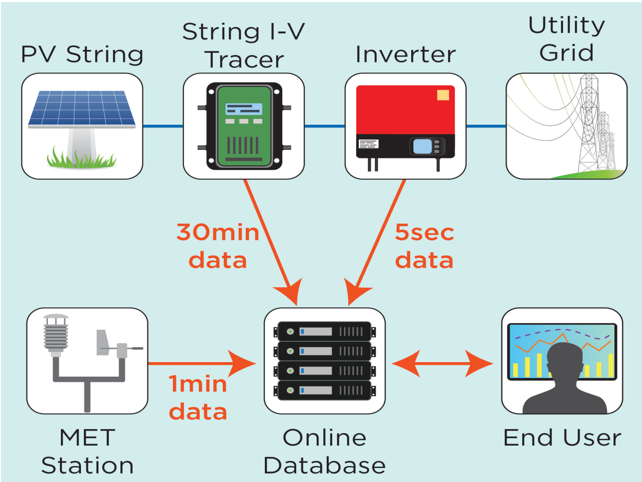


Figure 2: The addition of string-Level I-V curve tracers inline with each string enable remote intelligent diagnostics of PV faults in the cloud using machine learning techniques.

Data management & engineering

Proceeding from low to high-granularity monitoring levels means measuring, collecting, storing, retrieving, and analyzing larger amounts of data. Therefore, proper planning is needed to ensure the reliability of the data processing at all stages.

- Type of data: Two types of data are usually collected from the PV plants. The usual time series data consists of discrete measurements at each time-stamp (for example, inverter DC and AC parameters). However, I-V curve data is represented through a 2-D array of measurement values (of a few hundred or thousand points) at each time stamp. Therefore, the correct database is necessary to store and retrieve both time-series and I-V data. Time series data can be stored in any relational database management system (RDBMS), time-series database (TSDS), among other options. However, storing I-V data requires a relational database. Attempting to store HRMS data in a non-RDBMS server could lead to irretrievable data loss.
- Storing and Merging: Data collected from the field may not be integrated into the same database or database tables: inverter data may go to one server, and I-V curve data to another. However, ensuring that all data storing servers are synchronized to a common NTP (network time protocol) server clock or GPS clock is critical to the monitoring system. Otherwise, merging data from different servers with non-synchronized time stamps could lead to erroneous observations, results, analyses, and judgments.
- Frequency: IEC-61724 stipulates measurement frequencies based on the category of PV plants. This applies to conventional measurements like inverter parameters (time series data) and meteorological parameters. However, no standards exist for I-V data (I-V curves). Our work suggests a 40 – 60 min measurement frequency at medium to high irradiance is sufficient for measuring I-V curves. However, higher frequency measurement leads to more operational losses, and the I-V curve measuring equipment may wear unnecessarily.
- Filtering/Correction: Data filtering is essential to dropping unreliable data before performing analysis. Data collected during dynamic weather conditions, severe cloudy days, and due to faulty sensors is unreliable and needs to be detected and dropped. Some techniques, such as sky stability filters, daily per-unit filtering [3], or machine learning (ML) classifier (of normal and abnormal I-V curves) [4], can help filter inverter and I-V curve data. For example, a sky stability filter can effectively identify and drop dynamic weather data. For example, an I-V curve = measurement may be valid only if the sunlight conditions do not change more than a specified amount (~15 W/m2) for one minute before and after the I-V curve measurement. Additionally, comparing daily per-unit profiles of related

[3]. M. Matam et al, "Detecting Abnormal Profiles in the Database of a PV plant Through Programmatic Comparison of Per Unit Profiles," 47th IEEE PVSC, Calgary, OR, 2020, pp. 1534-1536.
[4]. M. Matam, et al, "An Algorithm for Filtering the Time-series I-V curves of a PV plant," 48th IEEE PVSC, June 20-25, 2021.

parameters (for example, inverter DC current, AC power, and POA irradiance) can help identify unreliable data. After performing data filtering, correcting the I-V curve data for irradiance and temperature will enable comparisons with previous years or peer data.

- Processing and Model Training: neural networks are powerful tools to model I-V data. We have implemented convolutional neural networks (CNNs) or long short-term memory (LSTM) neural networks [5,6]. It is a four-step process, and newer ML algorithms can leverage the existing approach. First, the irradiance and temperature corrected I-V curves are resampled with standard voltage array data. This helps to process I-V curves along the standard voltage array. Second, resampled I-V curves will be normalized to have a normal range of [0,1]. Third, NN predictors are calculated for all the I-V curves. Fourth, the model is trained and validated. The model considered in our project reported up to 96% detection accuracy at all irradiance conditions given the following PV faults: 1-module and 6-module partial soiling, cracked cells, cell interconnect breakage and increased series resistance in the string. Also, synthetic data can be used to augment the training dataset.

Why Consider Advanced Monitoring Early in the Design Process?

Integrating the monitoring equipment costs more at the operational stage than in the installation stage of the PV plant. Moreover, with advanced fault detection algorithms in place, it can detect defects in the PV modules at the infant stage of the operation, allowing the module replacement under applicable warranty claims and more yield.

[5]. M. W. Hopwood, et al, "Neural Network-Based Classification of String-Level IV Curves From Physically-Induced Failures of Photovoltaic Modules," in IEEE Access, vol. 8, pp. 161480-161487, 2020
[6] Hopwood, M.W.; Stein, J.S.; Braid, J.L.; Seigneur, H.P. Physics-Based Method for Generating Fully Synthetic IV Curve Training Datasets for Machine Learning Classification of PV Failures. Energies 2022, 15, 5085. https://doi.org/10.3390/en15145085

Diagnostic performance tradeoffs

The ability to detect and classify faults significantly differs at the inverter, combiner, string, and module levels. In addition, it varies based on the fault type and magnitude of power loss.

- Soiling is one of the most commonly occurring power loss events. The magnitude varies based on the plant's geographic location, climatic conditions (including precipitation and humidity), wind speed, and soil type. I-V curve data helps to know the soiling pattern across the PV array, the subsequent reduction in soiling due to precipitation, and assists in scheduling module cleaning. In addition, the data could help determine if PV array soiling is following a pattern, such as affecting mainly outer edge modules or modules in a specific outer edge (north, south, east, or west). Accordingly, cleaning can be directed to a particular section of PV modules and eliminate or reduce unnecessary cleaning activities. I-V curve data can accurately detect initial and severe soiling conditions. On the other hand, low-resolution inverter data cannot detect soiling unless it causes at least a 10% power loss.
- Cell cracks are pertinent problems for PV plants. They might occur in cell fabrication, module manufacturing, transportation, and installation. During PV plant operation, cell cracks could be caused by high-speed winds, hail storms, snow loading, and other weather-related stresses. In addition, cracks cause low power losses, making them impossible to detect using conventional inverter data. A characteristic I-V curve, collected at the module or string level, could help detect cracks after they have existed for some time. However, imaging techniques like electroluminescence (EL) and ultraviolet fluorescence (UVF) can detect cracks much earlier, but special equipment is needed to perform imaging, substantially increasing the cost.
- Cell interconnect failures usually occur at the cell edges, on the cells due to delamination, ribbon shunting, and string-interconnect failure. Prolonged

interconnect failure leads to cell hotspots and increased degradation rates. In some cases, interconnect failures could lead to arcing and fires based on the module's location in the string: modules at a potential further from the ground are more prone to arcing. Unfortunately, inverter data cannot detect these low power-loss faults, similar to cell cracks. However, after a sufficient time, I-V curve data of string or module can detect them.

- Module interconnect failure could occur in the module junction box or connectors. These are medium to high power loss causing faults that cannot be detected by the inverter data. However, the string I-V curve helps detect these faults, but not with the module I-V curve. These failures could cause a string to open-circuit, and thus reduce overall power generation. Open-circuited strings can also be detected with infrared (IR) imaging techniques.

Installation Considerations

The HRMS is not a thoroughly tested and commercialized concept. Therefore, a few challenges were anticipated in the design, purchasing, and installation stage. The following are a few challenges expected that need attention at the design stage of HRMS.

- Blocking diodes quality: A blocking diode at both positive and negative output terminals of each string prevents damage to the HRMS equipment when the string is disconnected/reconnected for I-V curve measurement. However, only a few commercial blocking diodes are available in the market, and their quality is questionable. Moreover, no standards are in place to stipulate the testing, evaluation, and certification of blocking diodes. In the absence of an international standard, an in-house mechanism is needed to test and validate the quality before installing them in the field.
- PV module MC connector issues: Module connectors are one of the challenges in installing the HRMS in an aged PV plant. Disconnecting the connectors of aged modules is complex; they may be welded after being in the field for years. Therefore, care should be taken not to cause damage.
- Meteorological sensors availability and location: Availability of irradiance and module temperature sensors is critical to the HRMS setup. Moreover, the sensor's must be installed as close as possible to the HRMS setup for the strings measured.
- Need for a powerful computer: Deploying CNN algorithms into the HRMS setup needs a powerful computer to initiate time-based I-V curve scanning, collect, process, store, and handle the data, and report important metrics to the PV plant operator. Remote access is critical to efficiently updating the CNN model.

Key Recommendations for HRMS setup

- *More care and planning are needed for setting up HRMS in an existing PV plant compared to a new PV plant.*
- *Carefully select HRMS hardware and confirm the integration of this new hardware is fire safe, secure, and reliable.*
- *Raise maintenance tickets as soon as HRMS raises flags. This ensures optimal benefits and achievable LCOE targets.*
- *Train the maintenance personnel on handling temporary and minor issues related to the HRMS setup.*
- *Perform additional diagnostic analysis and checks using the HRMS data.*

Cost-economics

Many factors must be considered in estimating the cost of designing, installing, and operating the HRMS setup.

- The PV plant voltage ratings and location impact the voltage rating of the hardware setup. Also, the higher the voltage, the more safety measures are included in the hardware, further increasing the cost.

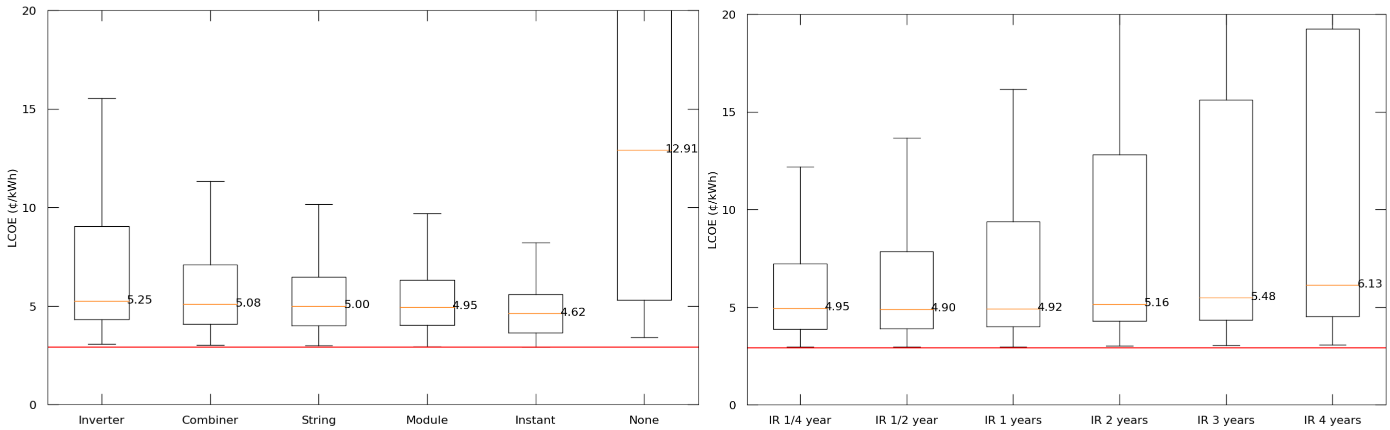


Figure 3: Comparing the LCOE (left figure) of the different monitoring methods. For example, the “Instant” LCOE shows the minimum LCOE when all the faults in the PV plant are cleared as soon as they occur, and “None” LCOE is a worst-case scenario which implies no fault detection and no response action. Aerial IR imaging-based LCOE analysis (right plot) is also presented. This is based on the premise that IR imaging detects all kinds of faults once they occur, but this is not the case in practice.

- The string's current rating impacts the switches considered for the string disconnection/reconnection operation. Therefore, a combination of electro-mechanical and semiconductor switches is suitable for high-voltage, high-current switching operations. In addition, it is important to select hardware that minimizes resistive power losses when in the nominal state (unswitched).
- The number of points on the I-V curve does impact the memory onboard the HRMS setup, processing speeds, and reporting. In addition, handling a high-density (higher points) I-V curve means a powerful processor needs to handle all stages, from initiating scanning to the end reporting, which increases the cost. Moreover, transferring the high-density HRMS data from a remotely located PV plant to the plant operator will also increase the data transfer charges.

LCOE benefits

Installing an HRMS setup is a cost to the PV plant operators. The validity of this cost depends on the ability of HRMS to detect faults in time, reduce losses, and incur more benefits. Our LCOE analysis confirms that HRMS is a potentially valuable asset in reducing the LCOE of the PV plant and adds economic value to the PV plant. The assumption here is that the HRMS operates reliably without any operational issues. Of course, the LCOE differs based on the plant's location, electricity purchasing agreements, etc. Also, in a relative comparison with the aerial IR imaging method, the widely accepted PV fault diagnostic approach is performed periodically or occasionally, yielding a comparable LCOE value. But it is on the premise that IR imaging accurately detects all types of PV faults, but this is not the case.

Download the Project details and reports

https://publications.energyresearch.ucf.edu/

Questions

Contact Hubert Seigneur, Project PI and Program Director, FSEC, UCF at seigneur@ucf.edu

UCF | FSEC Energy Research Center

UNIVERSITY OF CENTRAL FLORIDA