Reliability Assessment of Grid Connected Solar Inverters in 1.4 MW PV Plant from Anomalous Classified Real Field Data

Arif Sarwat[a], Patrick McCluskey[b], Sudip K. Mazumder[c], Miles Russell[d],
Sukanta Roy[a], Shahid Tufail[a], Shamini Dharmasena[a] and Alexander Stevenson[a]
[a]Department of Electrical and Computer Engineering, Florida International University
[b]Department of Mechanical Engineering, University of Maryland
[c]Department of Electrical and Computer Engineering, University of Illinois Chicago
[d]Director, Product Management, Yaskawa Solectria Solar
Email: [a]{asarwat,sroy007,stufa001,ikona001,astev028}@fiu.edu,
[b]mcclupa@umd.edu, [c]mazumder@uic.edu, [d]miles.russell@solectria.com

***Abstract*—In this work, a top-down analysis is carried out to investigate the impacts of environmental factors on the health, and hence on the reliability, of solar inverters (SI). Five years of real field data from 46 string inverters in a 1.4 MW Photovoltaic (PV) plant located at Florida International University (FIU) are used for the analysis. Collected data is classified and examined based on inverter faults, failures, and stress conditions using the classification and regression tree (CART) algorithm. Results have shown that inverter performance is highly correlated to ambient conditions, i.e. sunrise and sunset timing, relative humidity, and irradiance profile, and therefore adequate specific ventilation management can be a useful tool to mitigate some major inverter health issues. Triggered by this study, a prognostic analysis from the information in service tickets and machine learning (ML) outcomes will be carried out as future work.**

***Index Terms*—Solar Inverter Reliability, Real-field Data Analysis, Remote Condition Monitoring, CART algorithm**

I. Introduction

Renewable energy technologies, particularly solar energy, has a substantial impact on distribution grid characteristics due to their increased penetration. The failure of the solar inverter is one of the most prominent difficulties in the context of renewable energy system reliability. Thus, it is of high importance that the solar inverter, which controls real and reactive power injections of PV panels, be analysed for degradation tendencies. Overall reliability of SI is a complex problem to address since it involves large number of internal active and passive components and their individual reliability indices and failure mechanisms [1], [2].

To address this problem, different approaches have recently been taken which incorporate classification and machine learning based capabilities [3], [4]. The authors in [5] implement a systematic approach where ML text recognition is utilized to identify inverter related common failure modes. Additionally, because PV system data may be scattered, clustering algorithms can be particularly useful for deploying ML techniques [6]. In the work [7], decision tree algorithms have been investigated to classify PV power production under variable weather conditions and application of different state-of-art ML techniques on this regard is also compared. The authors in a sensor-based research effort, have explored a python-based library to detect anomalous events for grid-tied SIs [8].

In a similar work, an LSTM autoencoder technique was studied by the authors in [9] as a prognostic tool for the events in energy systems. On the other hand, environmental conditions are another highly varying factor for PV power systems. This is well addressed in [10], where the authors have predicted weather related interruptions from statistical data driven-models. Furthermore, a hybrid model to further improve PV generation estimation was developed and validated using several case studies [11] and power quality profiles [12]. Moreover, different architectures of commercially available PV monitoring systems, as described in [13], can be studied to get an idea on the working setup presented in this paper.

In this work, a large amount of reliability and environmental field data is mined using analytical methods. Next, field failures are classified into similar failure mechanisms and site groups based on prior field failure analysis with different environmental and operational conditions. Therefore, the objectives of this work lies in examining different inverter faults, classifying them based on defined criteria and performing CART algorithm based correlation study of key inverter parameters with weather data. The key contributions of the work are as follows: 1) finding evidences that SI incur different stress levels over a day's 24 hours time-span, 2) determine if stresses vary with solar irradiance variability and temperature rise, 3) inner stresses are transformed into electrical domain mostly in DC current carrier related alerts, events, and codes.

The rest of this paper is organized as follows: Section II outlines the FIU power system infrastructure and an overview of the data sources; Section III explains the methodology used for analysis and Section IV presents the key results and provides some discussions on their significance. Finally, Section V concludes the paper and provides details of future work.

II. Case Study of FIU Solar Plant

A. 1.4 MW Grid-Tied Solar Power Plant

The grid-connected 1.4 MW PV power plant at FIU has the capability to imitate mode-changing dynamics at the point of common coupling (PCC). The overall setup of the plant is presented by Fig. 1 and 2 [12], where the locations of recorders and data acquisition systems (DAS) are clearly shown. The data input to this work is collected from PQ recorders and DAS, which are operational in field from 2016 onward.

Fig. 1. Location of 46 grid-following solar inverter units

At a glance, Table I holds the setup and capacity information of FIU's plant. The accumulated AC voltage, AC current, harmonic distortion, flickers, frequency and power factors are logged in PQ recorder on 1 min timestamp interval. It also do continuous logging of significant event changes and abnormalities, which are translated to warning and fault messages. On the other hand, 46 units of three phase, grid-following string inverters' input-output data, codes, alerts are logged in DAS in similar 1 min time stamp. DAS is also responsible to deliver meter measurement and weather information on irradiance, temperature, and humidity percentages. These data are analyzed over the time and occurring frequency in this work.

B. Data Sources

The data from three main sources are analyzed in this work. Irradiance, ambient temperature, PV panel temperature, relative humidity, and lightning are the environment data considered for this study. These data points are logged by DAS. On other part, SI's input (DC voltage and current), output (3-phase AC voltage and current), and warnings, alerts and faults data of each individual inverter are used for the analysis. Further, the minimum, maximum, and average values of three phase output voltages and currents at the PCC and their flicker, total harmonic distortion (THD), and power factor data are taken from the PQ recorder. Next, service maintenance records consisting detailed failure analysis are considered into conclusiveness of deep learning based correlation study among environmental factors, alerts distribution and field failures.

Fig. 2. Revolution Wireless PQ Recorder location (LT sided) and DAS location for measuring time series data from inverters, meter and weather station and securely stores in a cloud server

III. Research Methodology

The SI of aforementioned 1.4 MW grid-tied solar power plant is considered for this case study. A unique combination of data analytic driven diagnostics and top-down system failure analysis have been used in this work. In the first part of it, long-term field performance data, inverter alerts are analyzed to assess the inverter stress levels. Therefore, environmental data which are found highly correlated, are analyzed on their role in creating stresses on inverters. Based on the alert distributions, weather variation and the chronological events of inverter field failures, an approach is taken to benchmark the inverters. The concluding part is attempted by iden-

TABLE I
1.4MW POWER PLANT SETUP LOCATED IN FIU

Total DC KW	Module Size (W)	Module Number	Capacity per Inverter (kW)	Inverter Type	Inverter Size (kW)	Inverter Number	Capacity Net Inverter (kW)	Capacity Net Module (kW)	DC/AC Ratio
1412.33	315.46	2x1520	25.22	3 Ph	24	6	144	151.321	1.05
	315.46	1x1340	31.525	3 Ph	24	40	960	1261.012	1.31

tifying some root causes and their possible mitigation plans.

IV. Results and Discussion

The research methodology laid out in section 3 is conducted and the following results were obtained. Initial subsections have demonstrated the analysis on inverters alerts distributions. Afterwards, a CART algorithm-based results are presented to identify the correlation between inverter DC voltage, current, environment impacts, and grid power which can be utilized to conduct exception detection leading towards performance fault estimation.

A. Historical Alerts From Inverters

The complete set of alerts from 2016 to 2022 are taken into consideration. This is for generating different insights such as maximum alert generating inverters, distribution of time between alerts, and time required to clear those alarms and codes. A set of communication related faults are excluded from analysis, with the fact that, there origin is not related to inverters rather to the network providers.

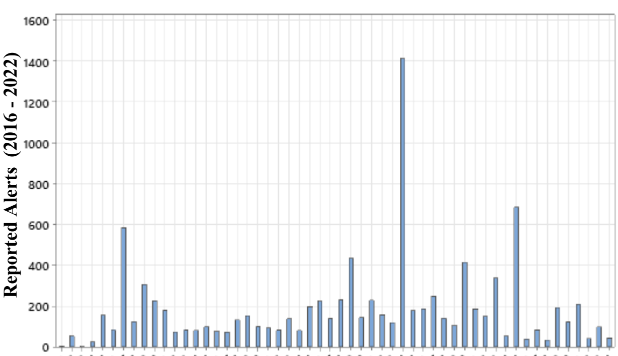

Fig. 3. Alert generation histogram per inverter

From Fig. 3, it is evidently that the top most erroneous inverter is inverter 34, followed by 45, 7, 29, and 40 in descending order. These 5 inverters will be in the spotlight of the further level of discussion in next steps. There are several alert names from historical data, whose distributions are shown in the Fig. 4.

The two most common alert sources are inverter and device communication alerts as shown in Fig. 4. The second highest failures originate from intermittent com-

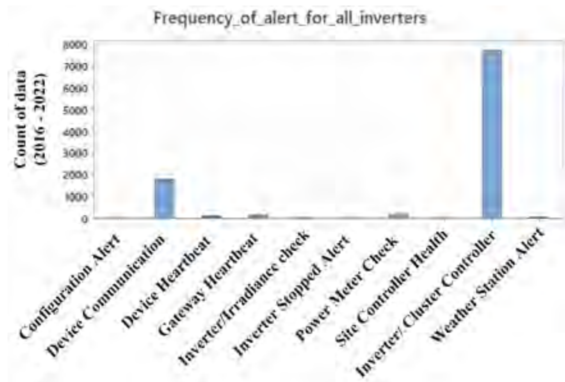

Fig. 4. Inverter alerts distribution histogram

munication failures and are excluded as stated earlier. The highest alerts are from inverters themselves, among which inverter 34 was the highest contributor as seen from Fig. 3. Its alert distribution is brought under study on this behalf and is presented in histogram plots of Fig. 5. The duration of faults and time before alert reset are plotted based on their frequency levels. The mean of power-related alerts of inverter no. 34 has a standard deviation as low as 1448. This high concentration of alarms are mostly DC attribute related, which is discussed in succeeding sections.

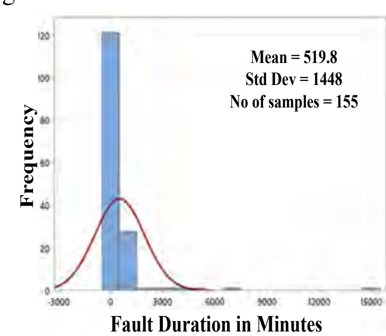

Fig. 5. Inverter 34 non-communication fault distribution (2016 to 2022)

B. Historical Alerts from Inverters' Dedicated Alert Server

Another alert server's data is analysed from a more recent time period of Dec 2021 to Mar 2022. The purpose is to add additional information from the inverter cluster controller which is collecting aggregated PV data from all 46 inverter units in real-time. The findings from this data can be seen in Fig. 6. The distribution pattern suggests that, interference device is the most common fault propagated from the 46 inverters in period of interest. The intended inverter's authorized troubleshooting manual has described this fault as a consequence of temperature rise issuisAccording to the manual [14], with interference device alerts, the inverter would stop due to excessive temperature and corrective measures such as to clean the fans, to ensure sufficient ventilation are suggested.

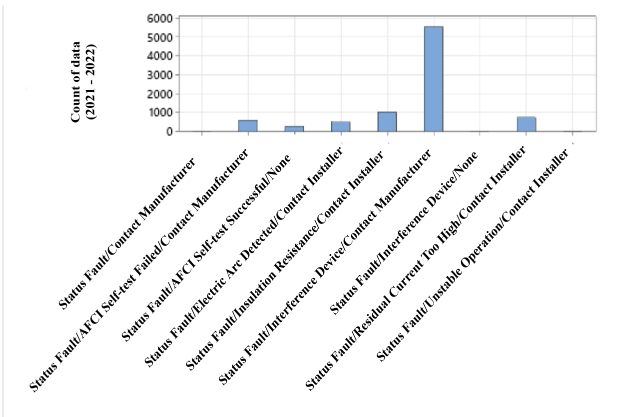

Fig. 6. Alert distribution from inverters' dedicated alert server

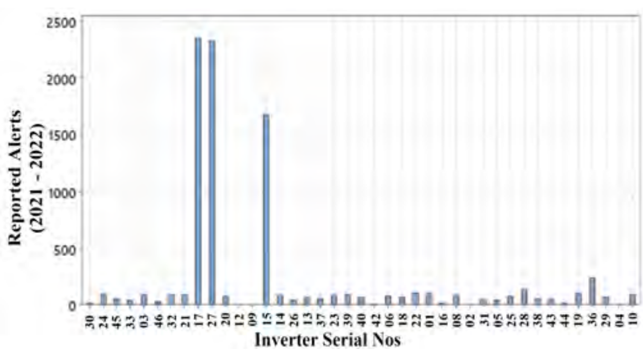

Fig. 7. Fault reporting inverters with fault counts (Dec, 2021 to Mar 2022)

In relating the inverters with the produced alerts, the results from Fig. 7 depicts that, the top three fault reporting inverters are inverter numbers 15, 17, and 27 . Among these three, inverter 17 was permanently failed as per the maintenance ticket information, showing that inverter 17 has a 6411 error code: 'Interference of Device'. Services reports from vendor described this code as related to lack of ventilation and temperature rise issues, where replacement is performed as permanent solution of it.

C. *Inverter Alerts on a Day (24 hours) Perspective*

The purpose of this sub-investigation is to observe the role that sunlight plays in creating stresses on the power circuit from the standpoint of alert/fault/warning timestamps over a 24 hour period. It is taken as an assumption that semiconductor devices, filters, and other elements which are comprising the core circuit of the inverters, produce thermal stresses when conducting current from a cold state [15]. In a similar way, stresses could be present during times of low irradiance where current flow may not be uniformly proportional as found in production power profiles as seen in Fig. 9.

Inverters' alerts server data is plotted per hour (average) over a 24 hour time span in Fig. 11 to investigate the highest stress time of the day using hourly increments. Around 7 A.M. the irradiance passes the threshold to activate all the Insulated Gate Bipolar Junction Transistors (IGBT) from a cold state, resulting in more alerts generated at this time. Most of the alerts are related to suddenly raised temperatures, electrical arcs, and insulation related problems. In the middle part of the day, alerts frequency is relatively stable, however when the sun sets, the stresses are shown to be increased on the inverter components resulting in higher amount of alerts around 7 P.M. of the day.

D. *Environmental Impacts and Correlation Study*

Stresses on inverters can be inferred from generated alarms or faults, as seen in Figs. 3, 4, 6, and 11. However, the source of these stresses is a resultant of environmental conditions to a great extent. Different aspects of the environment are considered, where weather patterns and seasonal changes as well as overnight moisture and dust concentrations are crucial factors to consider in this correlation study. Their relationship in how they impact PV production through inverters is complex to analyze and requires special techniques. In this work, inverter DC voltage correlation with environmental data is investigated by CART algorithm.

Results from the use of this technique are shown in Fig. 8 where the color concentration suggests that inverter DC input voltage has a negative correlation with humidity and very high positive correlation with irradiance. In Fig. 12, the top five correlating factors for inverter DC input voltage are extracted from the overall findings which are, inverter DC current, irradiance, module Temperature, ambient Temperature, and weather condition.

V. Conclusions and Future Work

In this work, an aggregated inverter alert list of six years is studied towards the goal of finding logical corre-

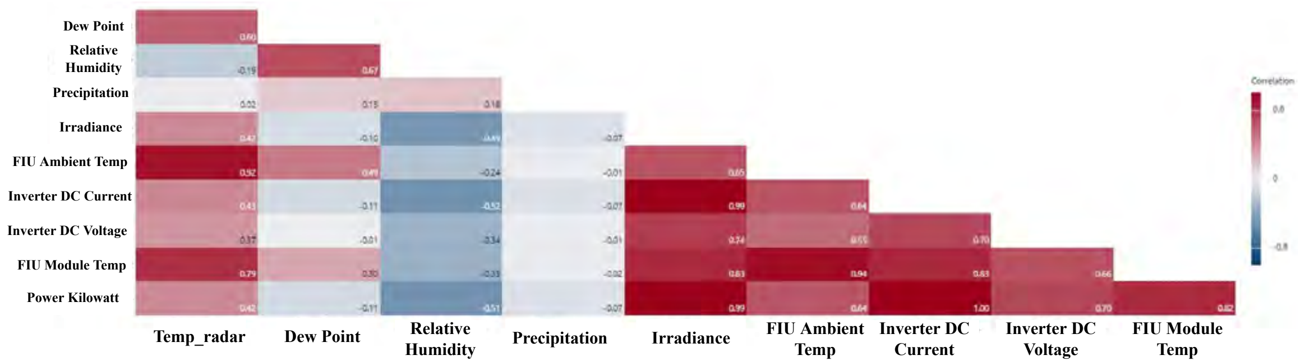

Fig. 8. Correlation study of inverter DC input with weather data

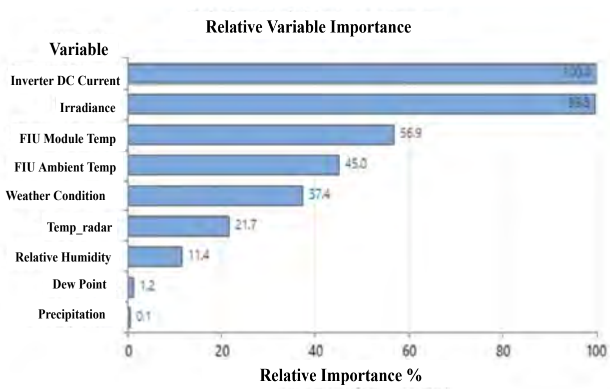

Fig. 9. Top five features in predicting inverter DC voltage

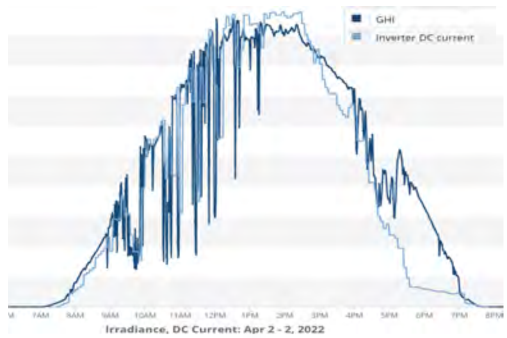

Fig. 10. Inverter DC input current with Irradiance profile over 24 hr.

lations with environmental data and inverter input-output characteristics. Depending on the configuration of the PV system, a single inverter's failure or reliability issues can cause a significant loss in total energy production. Thus, rigorous data point analysis and machine learning-based classification strategies have shown that, the inverter failures are highly correlated to stresses generated in specific periods of time during a day. Relative humidity and irradiance levels combined with their intermittent

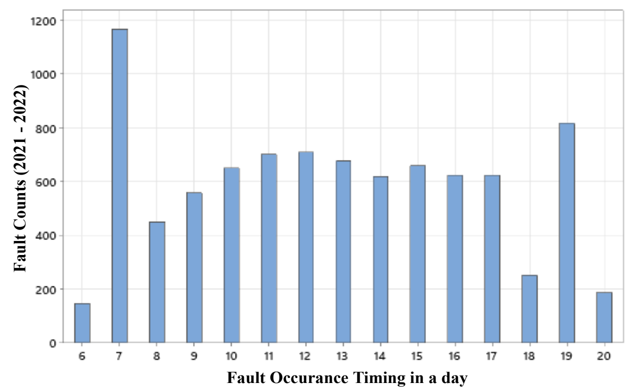

Fig. 11. Average fault reporting distribution over 24 hours of a day (Dec 2021 to Mar 2022)

behavior (and hence, changing ambient temperature) are key factors behind inverter warnings. Furthermore, the actual failure mechanisms have been found to be developed from DC current-related parameter warnings which is negatively impacting inverters' internal temperature rises. The findings are cross verified by the service records provided on the 46 inverter units. The field replacement of inverter no 17, and 34 and their failure analysis reports have been supportive to the conclusion that, physical relocation, scheduled ventilation, and dehumidification can be viable preventive measures against these failures The time shifting temperature rises also motivates the present research work for investigating the sun exposure of the outdoor inverters over the time as one of the next steps. This could aid the finding on the basis of the fact that, time bound ventilation will eventually decrease the stresses on component of inverters, hence increasing the lifetime of it. Lastly, this is a top-down approach, so more failure case studies are required to be done with the aid of unsupervised machine learning tools, which is a future scope. Eventually, component based stress and failure prediction model development is the broader part of this research, in securing the reliability of inverters.

ACKNOWLEDGMENTS

The authors would like to acknowledge that this material is based upon work supported by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technology Office (SETO) Award Number DE-EE-0009349. Furthermore, this report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

REFERENCES

[1] M. Dey, S. P. Rana, C. V. Simmons, and S. Dudley, "Solar farm voltage anomaly detection using high-resolution μpmu data-driven unsupervised machine learning," *Applied Energy*, vol. 303, p. 117656, 2021.

[2] L. B. Bosman, W. D. Leon-Salas, W. Hutzel, and E. A. Soto, "PV system predictive maintenance: Challenges, current approaches, and opportunities," *Energies*, vol. 13, no. 6, p. 1398, 2020.

[3] J. Ferrero Bermejo, J. F. Gomez Fernandez, F. Olivencia Polo, and A. Crespo Márquez, "A review of the use of artificial neural network models for energy and reliability prediction. a study of the solar pv, hydraulic and wind energy sources," *Applied Sciences*, vol. 9, no. 9, p. 1844, 2019.

[4] M. Ibrahim, A. Alsheikh, F. M. Awaysheh, and M. D. Alshehri, "Machine learning schemes for anomaly detection in solar power plants," *Energies*, vol. 15, no. 3, p. 1082, 2022.

[5] T. Gunda, S. Hackett, L. Kraus, C. Downs, R. Jones, C. McNalley, M. Bolen, and A. Walker, "A machine learning evaluation of maintenance records for common failure modes in PV inverters," *IEEE Access*, vol. 8, pp. 211 610–211 620, 2020.

[6] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Annals of Data Science*, vol. 2, no. 2, pp. 165–193, 2015.

[7] M. S. Nkambule, A. N. Hasan, A. Ali, J. Hong, and Z. W. Geem, "Comprehensive evaluation of machine learning MPPT algorithms for a pv system under different weather conditions," *Journal of Electrical Engineering & Technology*, vol. 16, no. 1, pp. 411–427, 2021.

[8] W. Dyamond and A. Rix, "Detecting anomalous events for a grid connected pv power plant using sensor data," in *2019 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA)*. IEEE, 2019, pp. 287–292.

[9] H.-S. Nam, C.-h. Shin, Y.-K. Jeong, and J. W. Park, "Fast anomaly detection model based on long sequences for energy storage systems," in *2021 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2021, pp. 1091–1094.

[10] A. I. Sarwat, M. Amini, A. Domijan, A. Damnjanovic, and F. Kaleem, "Weather-based interruption prediction in the smart grid utilizing chronological data," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 2, pp. 1–8, 2016.

[11] A. Sundararajan and A. I. Sarwat, "Hybrid data-model method to improve generation estimation and performance assessment of grid-tied pv: a case study," *IET Renewable Power Generation*, vol. 13, no. 13, pp. 2480–2490, 2019.

[12] A. Anzalchi, A. Sundararajan, A. Moghadasi, and A. Sarwat, "Power quality and voltage profile analyses of high penetration grid-tied photovoltaics: A case study," in *2017 IEEE Industry Applications Society Annual Meeting*. IEEE, 2017, pp. 1–8.

[13] M. M. Rahman, J. Selvaraj, N. Rahim, and M. Hasanuzzaman, "Global modern monitoring systems for pv based power generation: A review," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 4142–4158, 2018.

[14] "https://manuals.sma.de/sbxx-1av-41/en-us/43465739.html."

[15] N. R. Sorensen, E. V. Thomas, M. A. Quintana, S. Barkaszi, A. Rosenthal, Z. Zhang, and S. Kurtz, "Thermal study of inverter components," *IEEE journal of photovoltaics*, vol. 3, no. 2, pp. 807–813, 2013.