SAND2016-0794C

Wondering what to blame? Turn PV performance assessments into maintenance action items through the deployment of learning algorithms embedded in a Raspberry Pi device.

C. Birk Jones*, Manel Martínez-Ramón†, Olga Lavrova*,
Craig Carmignani*, Joshua S. Stein*, and Bruce H. King*

*Sandia National Laboratories Solar PV & Grid Integration, Albuquerque, NM, U.S.A
†Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, U.S.A

***Abstract*—Advanced monitoring of photovoltaic (PV) systems can insure efficient operations. However, extensive monitoring of large quantities of data can be cumbersome. The present work introduces a simplified, cheap, yet effective data monitoring strategy for classifying behavior, and determining lost revenues automatically. This was achieved through the deployment of Raspberry Pi (RPI) device at a PV system's combiner box. The RPI was programmed to collect PV data through Modbus communications, and store the data locally on a MySQL database. Then, using a Gaussian Process Regression algorithm the RPI device was able to accurately estimate string level current, voltage, and power values. Based on the residuals between this estimate and the actual values the RPI was able to classify the behavior as normal or in a fault condition using a Laterally Primed Adaptive Resonance Theory neural network. In addition identifying the fault condition, the RPI output the potential lost revenue caused by the abnormal condition. The information produced by the RPI could help define maintenance activities in real-time so that problems could be addressed and solved quickly.**

***Index Terms*—fault classification, Gaussian Process, Laterally Primed Adaptive Resonance Theory, Raspberry Pi**

I. Introduction

Solar photovoltaic (PV) arrays require minimal maintenance and operations to produce electricity. This is true, because fixed tilt arrays have minimal moving parts. Whereas, gas generators or a wind turbines have components that must be oiled, repaired, or replaced on a regular basis. Therefore, common practice for many systems has been to install the PV panels and then walk away. However, faults do occur, and financial investments in large PV arrays demand strict performances tolerances so that returns match expectations. Therefore, large scale developments require intensive monitoring and oversight.

Current recommendations for PV system oversight suggest metrics such as performance ratio (PR) [1], temperature corrected PR, Energy Performance Index (EPI) SAM model, EPI Regression model, and Power Performance Index (PPI). However, these approaches do not provide an effective means to detect and classify faults. Instead, the present work proposes the implementation of a low cost intelligent Raspberry Pi (RPI) device. The device can be deployed at combiner boxes or inverters to collect actual data through Modbus communications. Then, the PV sub-system data can be analyzed by the onboard analytics that includes advanced machine learning algorithms.

The integration of an intelligent RPI devices into a PV system can perform automatic data collection and assessments that can lead to improved operations. This paper introduces a methodology where the RPI computes an estimate of system performance using a Gaussian Process Regression (GPR) algorithm. The estimated values were than compared with the actual, and the differences were assessed by the Laterally Primed Adaptive Resonance Theory (LAPART) neural network to detect and classify sub-system faults. The approach not only classifies system anomalies but it also uses the estimated performance to compute a lost revenue caused by the fault condition. This abstract describes the implementation of the RPI monitoring system and the embedded algorithms in the methodology (Section II). It then outlines key results and provides concluding remarks in Sections III and IV respectively.

II. Methodology

The present work proposes the integration of an intelligent, cheap, and deployable RPI device into a PV system that can detect and diagnose fault conditions. Machine learning algorithms, that include GPR and LAPART, were used to estimate actual performance and then classify the status of a PV system as shown in Figure 1. The process began with the extraction of weather and actual system data through modbus communications. The weather data was presented to the GP estimator to define the ideal current, voltage, and power values.

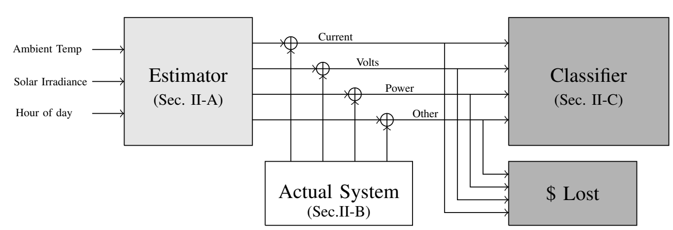

Fig. 1. The intelligent RPI included an embedded platform that first estimated performance using a gaussian process algorithm. It then calculated the difference between the actual and the estimate. The patterns in the differences were then evaluated by a second learning algorithm that detected and diagnose fault conditions. Finally, the difference between the actual and estimated values were used to compute the lost revenue caused by the fault condition.

The estimated values were then subtracted from the actual values produced by the system. These residual values were used by the classification algorithm to detect and diagnose fault conditions. In the event of a fault, the calculated residuals were used to estimate the amount of money lost due to the abnormal condition.

A. Estimate PV Performance

The estimated PV performance, described as the *Estimator* in Figure 1, can be calculated using a component-based or empirical model. For example, the python version of PV_LIB (https://github.com/pvlib/pvlib-python) could be run on a RPI device. However, in this experiment the estimator was the GPR algorithm. The algorithms were presented with a training data set, $D = (x_i, y)|i = 1, ...n)$. The inputs x_i included ambient temperature, solar irradiance, and hour of the day, and outputs y_i were current, voltage, and power.

GP can be defined as a set of random variables where any finite number of the set have a joint Gaussian distribution [3]. GP applies a distribution over functions that are specified by a mean function and a covariance function as shown in Equation 1.

$$f(\mathbf{x}) \sim GP(\mu(x), k(x, x')) \tag{1}$$

The mean function, $\mu(x)$, is usually defined to be zero and the covariance $k(x, x')$ defines the prior properties of the functions considered for inference [4]. The k in the covariance represents the kernel function which projects the data into a higher dimensional feature space to increase the computational power of the algorithm [5].

B. Actual PV System

The intelligent RPI was deployed to monitor an actual 10.8kW$_e$ system (Figure 2) located in Albuquerque, New Mexico. The array had four strings of 10 modules that were combined into into one prior to entering the inverter. The

Fig. 2. The present work performed tests on a 10.8kW$_e$ array. The array has four strings that each have 10 modules. The strings are combined prior to entering the inverter. In addition, each of the string's current and voltage are monitored.

RPI device, shown in Figure 3, was deployed inside the combiner box and offered an energy efficient way to perform computational tasks [9]. A critical task was collecting actual sensor data which was achieved through Modbus communcations. The device connected with the Modbus TCP/IP network created by the Gantner Q.Station 101 data collection device through the ethernet port. A Modbus serial connection was achieved through the RPI USB port, which provided access to inverter data. The data was extracted from the two Modbus protocols using Python programing language script and stored in a MySQL database on the RPI.

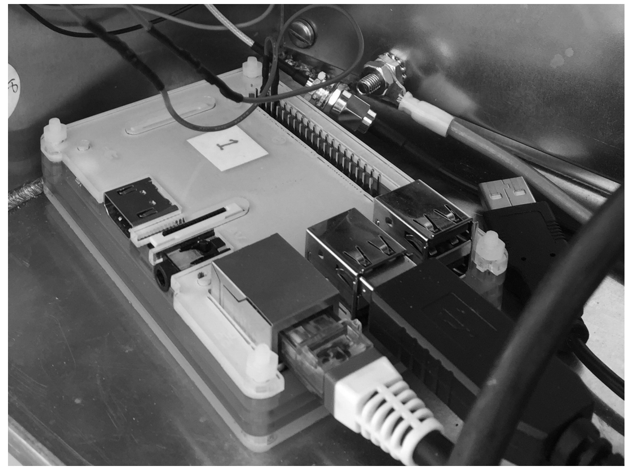

Fig. 3. The Raspberry Pi was located inside an enclosure attached to the array. It was powered by a 5V power supply, and connected to the existing data collection devices through Modbus TCP/IP and Serial ports. It also has its own GPS time clock to maintain correct date and time throughout the data collection process.

C. Classification of Faults

The classification of faults was performed using a LAPART algorithm. The algorithm was introduced by Healy and Caudell for logical inference and supervised learning [6]. The LAPART algorithm can converge rapidly towards a clear solution because it does not depend on the gradient descent method that is used in many popular algorithms such as the multi-layer perceptron. The gradient descent approach is susceptible to issues that include slow and/or incorrect convergence to the optimal solution [7]. The LAPART architecture couples two Fuzzy Adaptive Resonance Theory (ART) algorithms to create a mechanism for making predictions based on learned associations.

III. Results

The results for this experiment were broken out into three sections. The first section (Sec. III-A) describes the ability of the GPR to estimate PV performance.

A. PV Performance Estimates

The GPR algorithm estimated current for each string, overall voltage, and total power. The estimation could be performed in real-time or at the end of each day. GPR was able to estimate ideal performance well. For example, Figure 4 plots the actual and estimated values for a single day. The results of the classification of the residual values performed by the LAPART algorithm were documented in Sec. III-B.

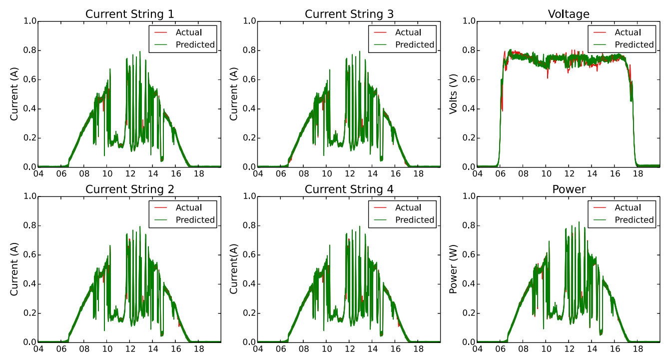

Fig. 4. The estimated PV performance for current, voltage, and power fit well with the actual values as shown for this single day of intermittent behavior.

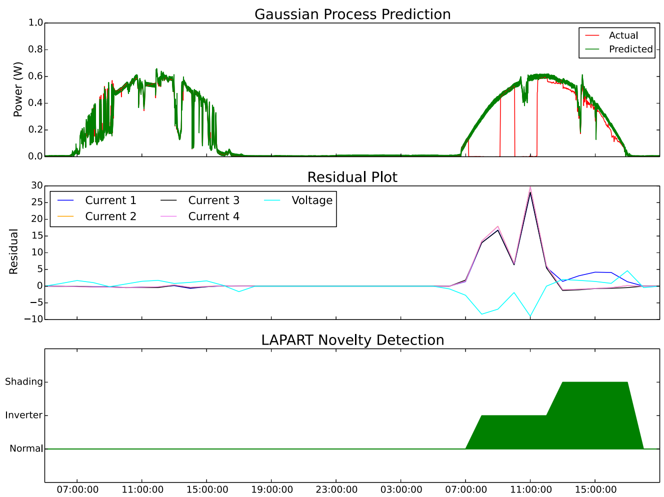

Fig. 5. Top: Actual and estimated power for a two day period. Middle: The residual current and voltage indicates significant faults on the second day. Bottom: The LAPART algorithm was able to classify the fault conditions

B. Fault Classification

The fault classification process used the LAPART algorithm and could accurately classify normal, inverter, and shading characteristics. For instance, two fault conditions were introduced into the actual system. The first step, which was to perform an estimate of the performance using the GP algorithm, was conducted. The actual and estimated results are plotted in the top graph of Figure 5. Then the difference between the actual and estimated were computed and plotted in the middle of Figure 5. Finally, the LAPART algorithm considered the residual patterns and was able to identify the normal and two fault conditions (inverter shutdown and shading) over the course of a single day as shown in the bottom graph of Figure 5.

C. Lost Production

The final task, performed by the intelligent RPI, was to estimate the lost production caused by the fault conditions. The difference between the GP estimate and the actual power results were calculated and described in Figure 6. In this case, the faults caused the overall energy production to reduce by 22.9kWh for the 10.8kW array.

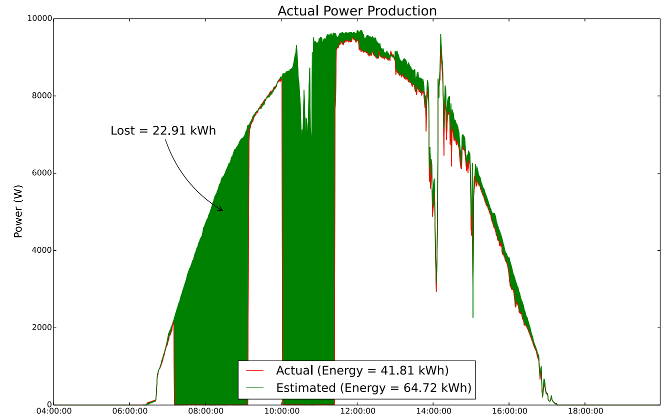

Fig. 6. The difference between the actual and estimated was used to calculate the lost energy. In this case a total of 22.9kWh were not produced.

IV. Conclusion

The RPI device was integrated into an existing PV array and successfully collected data through Modbus communications. The GP algorithm, embedded in the RPI, was able to accurately estimate current, voltage, and power values. The classification of faults was achieved using the LAPART algorithm. Finally, the full paper will include a more detailed review of estimate and classification accuracy for multiple fault conditions.

Acknowledgment

This work was supported by the U.S. Department of Energy SunShot Initiative

Sandia National Laboratories is a multi-program managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

References

[1] V. Sharma, A. Kumar, O. S. Sastry, and S. S. Chandel, "Performance assessment of different solar photovoltaic technologies under similar outdoor conditions," *Energy*, vol. 58, pp. 511–518, Sep. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544213004994

[2] J. Mokri and J. Cunningham, "PV System Performance Assessment," SunSpec Alliance, Tech. Rep. 2, Jun. 2014.

[3] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT Press, Aug. 2012.

[4] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. University Press Group Limited, Jan. 2006.

[5] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, Mar. 2000.

[6] M. Healy, T. Caudell, and S. Smith, "A neural architecture for pattern sequence verification through inferencing," *IEEE Transactions on Neural Networks*, vol. 4, no. 1, pp. 9–20, Jan. 1993.

[7] S. Haykin, *Neural Networks A Comprehensive Foundation*, 2nd ed. Prentic-Hall, Inc, 1999.

[8] "Raspberry Pi," Nov. 2015. [Online]. Available: https://www.raspberrypi.org/

[9] E. Upton and G. Halfacree, *Raspberry Pi User Guide*. John Wiley & Sons, Sep. 2014.