A Hybrid Regression Algorithm for Solar Power Forecasting

Kathy F Turner, Sahuaro High School, Kristen Jaskie and Andreas Spanias . SenSIP Center, School of ECEE, ASU

***Abstract* - Solar energy forecasting has been an open problem for power utility companies. The central problem is anticipating the contribution from solar fields and from individual rooftop installations. Using previous work as a foundation, we develop an algorithm to predict the energy production during different times of day. More specifically, we will use 12 months of data from a single residential solar array and use our algorithm to predict future output of the PV array. The data was recorded in 15-minute intervals from June 2019 to June 2020.**

***Index Terms*—PU learning, Machine Learning, Solar Energy, PV Forecasting, Regression**

PROJECT DESCRIPTION

Accurate Energy forecasting with fixed and renewable energy sources is an open problem for utility companies and grid control. Prior work [1,2] has shown that the use of a vector autoregression framework for solar power forecasting has been successful. Existing research [3-6] in forecasting solar energy production has focused on large Photovoltaic (PV) arrays for multi-residential, business, and utility scale plants. Predicting the energy production for small solar arrays single family residences has not been extensively studied. The objective of this RET project is to study machine learning (ML) algorithms [7,8] for use in energy forecasting. The research is part of SenSIP workforce programs [9] which include load forecasting [10,11] solar monitoring and control [12-14]. Identifying energy production during different times of the day can be used to determine the mean energy production based upon the weather for a particular day. Given this data, a user will have the information to determine what energy production on future dates. Residential users will be able to predict the load needed in a battery backup, or usage from the main power grid.

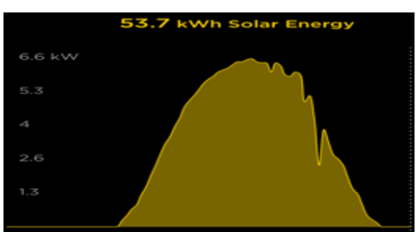

Figure 1: f Real-time solar energy production measured for a Tucson residential property.

A sample solar energy profile for residential property is shown in Fig. 1. The years' worth of data provides the information to train and a test of the forecast algorithm. We plan to use several ML algorithms and use fusion to improve forecast results. More specifically, our research will explore ML methods for solar power forecasting with the emphasis on regression methods including linear regression, logistic regression, and multivariate regression. Our approach will first consider synthetic data to help create and validate the algorithms. Further study will document the development of a customized regression algorithm which will be profiled in terms of performance and complexity. Anticipated outcomes include: a) algorithm performance characterization, b) python code implementation, and c) comparison of the algorithm performance with real and synthetic data. Positive Unlabeled (PU) Learning [15] methods will also be explored.

References

[1] Assouline, D., Mohajeri, N., & Scartezzini, J. L. (2017). Quantifying rooftop photovoltaic solar energy potential: A machine learning approach. *Solar Energy*, *141*, 278-296.

[2] Lonij, V. P., Jayadevan, V. T., Brooks, A. E., Rodriguez, J. J., Koch, K., Leuthold, M., & Cronin, A. D. (2012, June). Forecasts of PV power output using power measurements of 80 residential PV installs. In *2012 38th IEEE Photovoltaic Specialists Conference* (pp. 003300-003305).

[3] A. Tuohy et al., "Solar Forecasting: Methods, Challenges, and Performance," in IEEE Power and Energy Magazine, vol. 13, no. 6, pp. 50-59, Nov.-Dec. 2015,

[4] Fumo, N. and Rafe Biswas, M., 2015. Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, 47, pp.332-343.

[5] H. Yang, C. Huang, Y. Huang and Y. Pai, "A Weather-Based Hybrid Method for 1-Day Ahead Hourly Forecasting of PV Power Output," in IEEE Transactions on Sustainable Energy, vol. 5, no. 3, pp. 917-926, July 2014,

[6] J. Huang, Y. Liang, H. Bian and X. Wang, "Using Cluster Analysis and Least Square Support Vector Machine to Predicting Power Demand for the Next-Day," in IEEE Access, vol. 7, pp. 82681-82692, 2019

[7] U. Shanthamallu, A. Spanias, C. Tepedelenlioglu, M. Stanley, "A Brief Survey of Machine Learning Methods and their Sensor and IoT Applications," *Proc. 8th IISA 201),* Larnaca, August 2017.

[8] K. Jaskie and A. Spanias, "Positive and Unlabeled Learning Algorithms and Applications: A Survey," *Proc. IEEE IISA 2019*, Patras, July 2019.

[9] A. Spanias, Machine Learning Workforce Development Programs on Health and COVID-19, *Proc. IEEE IISA 2020*, Piraeus, July 2020.

[10] D. Smith, K. Jaskie, J.Cadigan, J. Marvin, and A. Spanias, "Machine Learning for Fast Short-Term Energy Load Forecasting," *IEEE ICPS 2020*, Tampere, June 2020.

[11] K. Jaskie, D. Smith, A. Spanias, Deep Learning Networks For Vectorized Energy Load Forecasting, Proc. *IEEE IISA 2020*, Piraeus, July 2020.

[12] H. Braun, S. T. Buddha, V. Krishnan, A. Spanias, C. Tepedelenlioglu, T. Takehara, S. Takada, T. Yeider, and M. Banavar, Signal Processing for Solar Array Monitoring, Fault Detection, and Optimization, Synthesis Lect. Power Electronics, J. Hudgins, Ed. Morgan & Claypool, vol. 3, Sep. 2012.

[13] A. Spanias, "Solar Energy Management as an Internet of Things (IoT) Application," *Proceedings 8th IEEE IISA 2017*, Larnaca, August 2017.

[14] S. Rao, S. Katoch, V. Narayanaswamy, G. Muniraju, C. Tepedelenlioglu, A. Spanias, P. Turaga, R. Ayyanar, and D.Srinivasan, "Machine Learning for Solar Array Monitoring, Optimization, and Control,"Synth. Lect. Power Electronics, Morgan & Claypool, Ed. J. Hudgins, 91 pages, , Aug. 2020.

[15] K. Jaskie, C. Elkan, A.Spanias, A Modified Logistic Regression For Positive and Unlabeled Learning, *IEEE Asilomar Conf. Sign., Syst. & Comp.*, Monterrey, 2019.

Acknowledgment: The RET is supported by NSF Award 1953745