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: 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

Acknowledgment: The RET is supported by NSF Award 1953745