NASA Ames Summer 2013 Progress

Dan O'Leary & Samuel Kahn

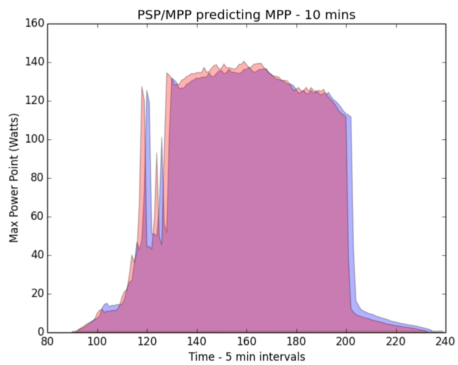
September 21, 2013

The two main goals we set out to complete at the start of the summer was to begin storing data from the NASA Ames Renewable Energy Testbed and then use this data to make predictions about future power production. We decided to store the data on the MySQL server at UCSC. A Raspberry Pi is set-up to run our Python programs, which pulls data from the sensors, temporarily stores it on a local SQLite database on the Raspberry Pi and then pushes it to the MySQL database at UCSC.

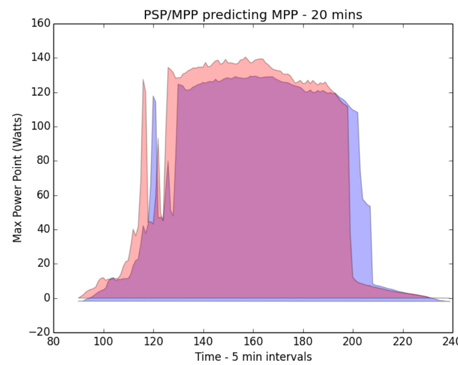
Currently, we have a cron job automating this process and sending data to the MySQL database at UCSC on 5 minute intervals. We are currently storing data for the precision spectral pyranometer, normal incidence pyroheliometer, and max power point. We have been unable to maintain a steady connection via bluetooth with the orientometer, but once this is resolved we are in place to begin storing azimuth and altitude readings from the solar array. In addition, the schema we have designed will be used on all of the databases for the relevant renewable energy test beds.

We have also developed various Artificial Neural Networks using the PyBrain Machine Learning Library allowing us to test various different neural network architectures. Our current architecture consists of either one, two or three input nodes, 20 hidden nodes in our hidden layer and one output node, which is our prediction for the max power point. The high number of hidden nodes seems to be better for fitting complicated functions, of course this is an area where we need to do more testing. The below figures show the our results for predicting the max power point 10 minutes, 20 minutes, 30 minutes, and 60 minutes in advance. Figure 1 and 2 shows a single day of data and twelve days of data respectively. The max power point and precision spectral pyranometer are used as our feature vector and we are prediction the max power point in the future.

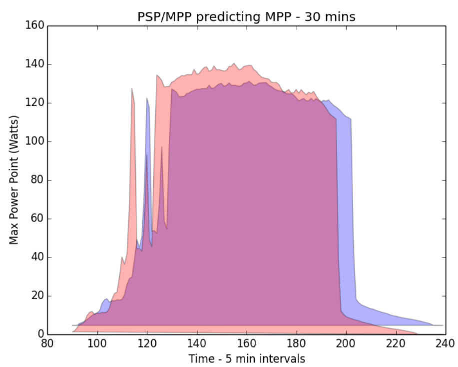
We would like to improve our persistent model to be able to predict dawn and dusk better (we found that dawn and dusk have systematically the highest prediction error). Also, we need to add error detection so that we are using a data without any temporal gaps in it — this causes error and poor results. Our network architecture still needs to be optimized and texted with a suite of different type of predicting feature vectors, hidden nodes and transfer functions. There is also literature stating the Support Vector Machines are excellent for predictions of these natures, so in the near future we should begin exploring other machine learning algorithms



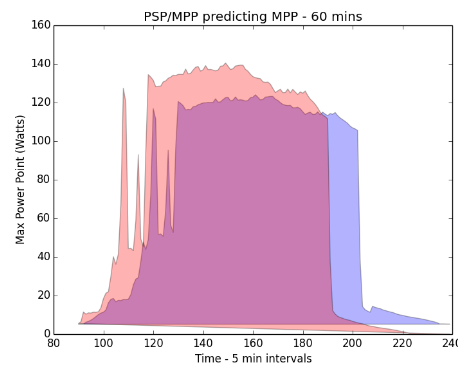
(a) 10 minutes



(b) 20 minutes

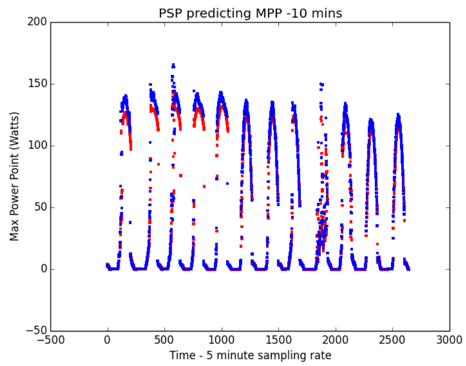


(c) 30 minutes

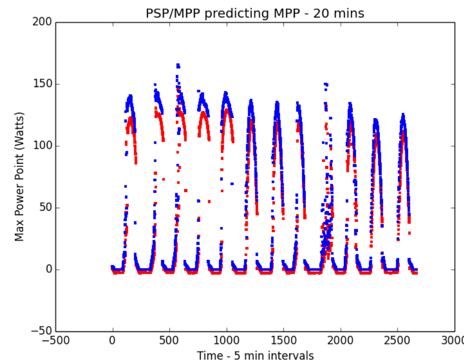


(d) 60 minutes

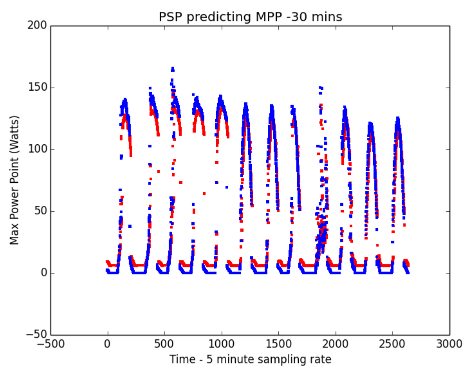
Figure 1: Predictions for single day at intervals of 10 minutes, 20 minutes, 30 minutes and 60 minutes using Max Power Point and Precision Spectral Pyranometer to predict Max Power Point where x-axis is time (5 minute intervals) and y-axis is value of max power point. Blue is prediction, pink is real value.



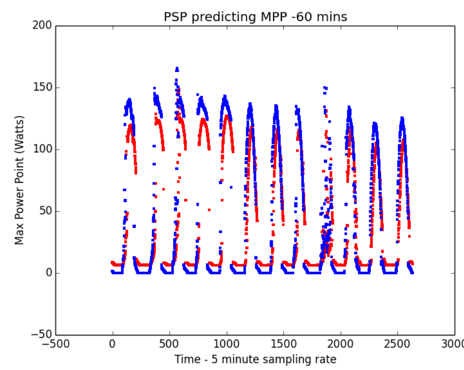
(a) 10 minutes



(b) 20 minutes



(c) 30 minutes



(d) 60 minutes

Figure 2: Predictions over twelve days at intervals of 10 minutes, 20 minutes, 30 minutes and 60 minutes using Max Power Point and Precision Spectral Pyranometer to predict Max Power Point where x-axis is time (5 minute intervals) and y-axis is value of max power point. Red is prediction, blue is measured valued.