Energy Disaggregation

Goal

Extract Energy consumption time-series for the Pool-Pump from the time series energy consumption data for a home over 1 year period sampled at 15 min granularity.

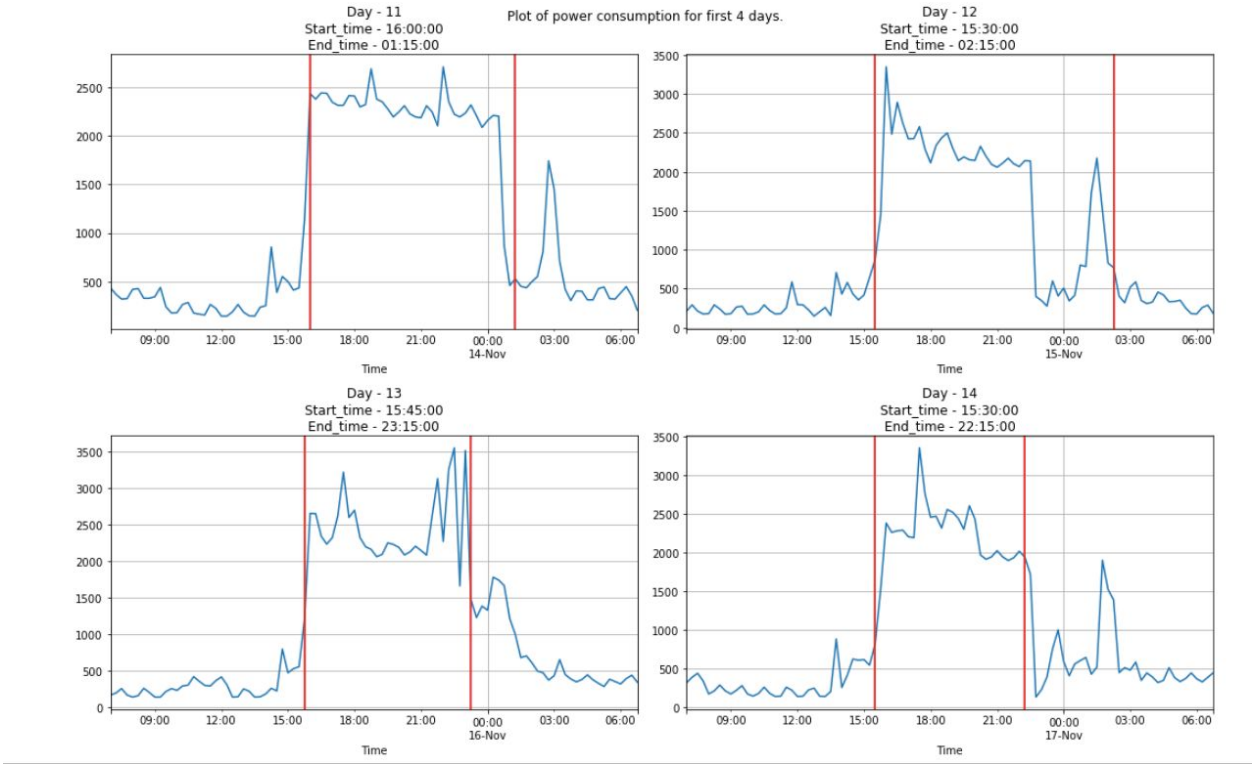
This is aggregate energy consumption data for one home that contains different appliances including HVAC, water heater, pool pump, etc.

Assumptions

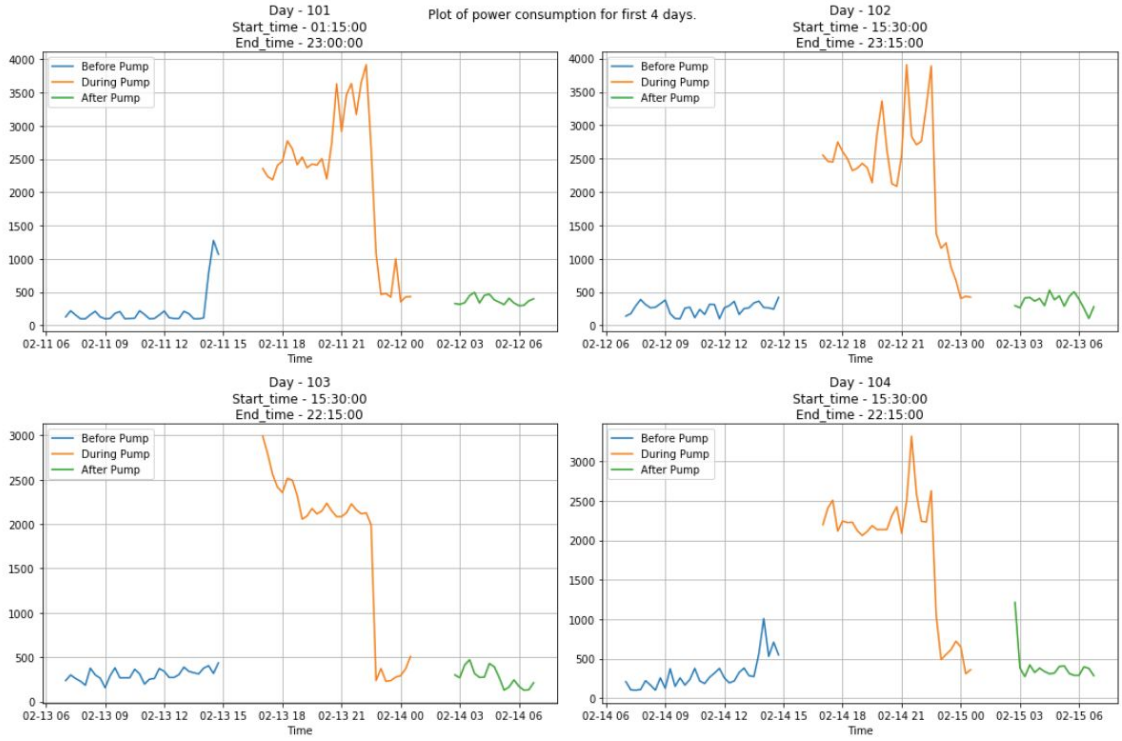
1. The Pool Pump starts and stops at fixed timings everyday.
2. It maintains consistent power consumption during its operations.
3. Its Power rating is between 1000W to 3000W.

Techniques used

1. The Energy data in Wh is loaded into the python environment as a dataframe using the pandas package. It is then converted to Power-load (in Watts) series by multiplying by 4, as Energy at 15 mins -> Power/4.
2. To Smoothen out the noise, A rolling average of dimension 5 is taken over the Power data.
3. Now the 1st difference is taken to extract the timings of maximum increment and decrement in the power-load
4. Now the start-time and end-time of the pump is estimated to be when there is maximum power-delta in the +ve and -ve directions respectively.

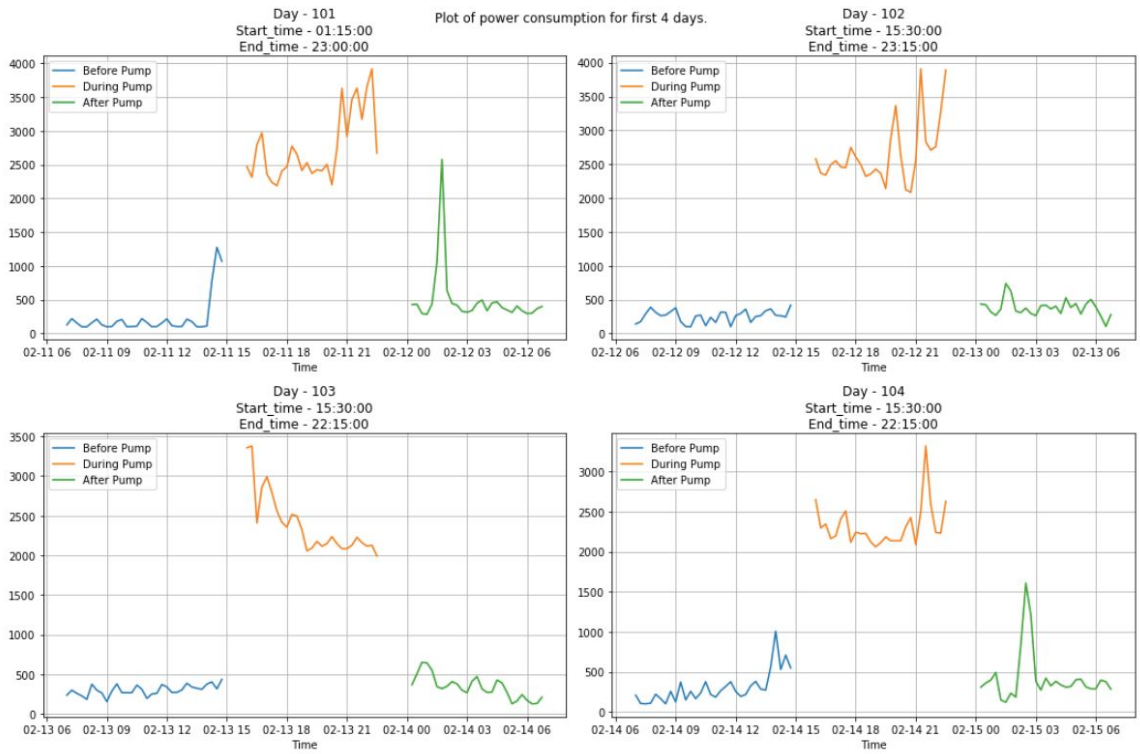


Visualizing Power-data vis-a-vis estimated start-time and end-time of the pump



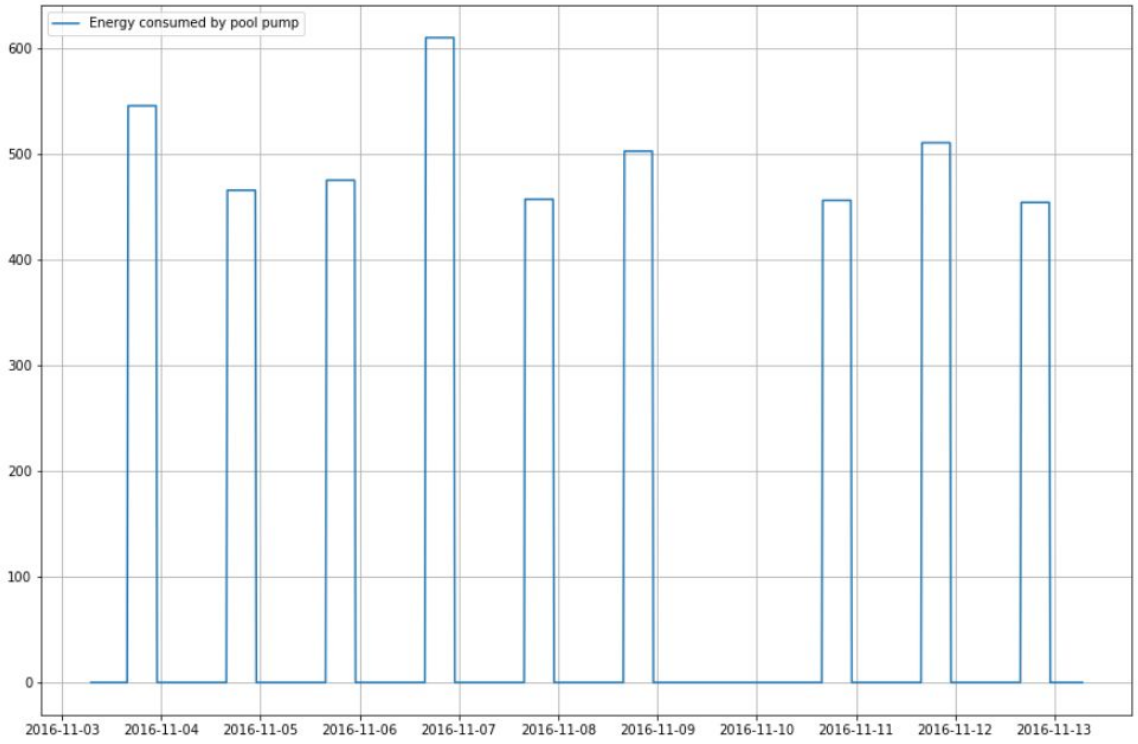
Visualizing power-load patter - before Pump operation, during and after the pump is stopped.

5. From start and end timings of all the days, based on the majority voting (highest frequency), choose global start and end timing of the pump across all days.



Visualizing power-load patter - before Pump operation, during and after the pump is stopped, at the global start and stop timings, calculated above

6. Remove the power values before and after the estimated pump operational timings , from the main series, by making those as zero.
7. Remove power-load of other appliances from the total power during (estimated) pump operational timings, by subtracting median power during off-pump duration from the on-duration.
8. Now we have estimated the power-load of the pump for each day. We set its lower-limit as maximum of 1000W and 1st Quartile value. Upper-limit as minimum of 3000W and 3rd Quartile. For any day, if the estimated power-load is less than the lower-limit, we reduce it to zero, assuming the pump wasn't used that day. If it's above the upper limit, we clip it to the upper-limit. Else the value stays the same.
9. Finally, we divide the power series by 4 to get the energy consumption series in Wh.



Visualize first 10-day energy consumption of the pump

Time Series output of pool pump consumption

Please refer to the **Pump_energy_series.csv** file submitted.

Insights acquired while working on this data

- Gained perspective on basic energy disaggregation problem.
- Hands-on experience with time-series analysis and tools.
- Proposing hypothesis and implementing a solution based on that.

Interesting visualizations / observations

Please refer the jupyter notebook -
Energy_Disaggregation_Vishal_Submission_notebook.pdf /
Energy_Disaggregation_Vishal_Submission.ipynb

The computational complexity of your technique(s)

O(# data points in 1 day). As the code made use of matrix/vector calculation advantages.