Grid Services NRE Detection – Load Shift

Analysis Methodology

The first script, ‘0. Load Packages’ lists all the packages required for the analysis. Please install any that are not already available on your system. The items below provide details on all other scripts used in the analysis.

1. **Load & Process Data:** In this work, we have analyzed three modeling algorithms’ ability to predict energy load under the conditions of baseline degradation. The predictions are done over the weekdays in April 2014, and the baselines consist of weekdays only as well. The baselines undergo degradation, ranging from 0 days of degradation to 50 weekdays of degradation.

Due to this setup, the baseline period can extend back to October 2013: to predict 04/01/2014, a 70-day baseline for TOWT with only weekdays and a 50-weekday baseline degradation starts in 10/15/2013. Note that the 50-weekday degradation and 70-weekday baseline only excludes weekends and not public holidays. Excluding public holidays may be considered and included as a step in the future.

In order to extend the data to October 2013, the pre- and post-datasets were combined. Next, only datasets that had at most 10% of the data missing in the time period October 2013 – April 2014 were considered in the analysis.

1. **Apply TOWT and Apply Day-Matching:** These two scripts apply three algorithms (TOWT with a 7-weekday baseline, TOWT with 70-weekday baseline, and Day-Matching with 10-weekday baseline) for each meter, each prediction hour of the weekdays in April 2014, with 0 through 50 weekdays of baseline degradation. The results are stored as data frames in the .rds file format in the ‘TOWT Interim Data’, ‘TOWT Interim Data – 70 day baseline’, ‘DM Interim Data’ folders, respectively. In order to be consistent with the amount of missing data rule, the baseline models are constructed, and predictions are made if no more than 90% of the data in the baseline period is missing. Predictions are only made for April 2014 weekdays that have all 24 hours of data available.

Note that the fractional missing data rule has a significant impact on the aggregated results. The permissible fraction of missing data is an input in the analysis and can be changed as needed.

1. **Calculate NMBE:** This script calculates residuals for each prediction, for each meter and each hour, and also calculates groupwise NMBE[[1]](#footnote-1): NMBE calculated over all predictions (all hours of the day: 0:00 – 23:00) for each meter. So, the result of this step is a list of 51 data frames (representing 0 through 50 weekday degradation), with each dataframe holding the NMBE values for each meter, considering all predictions over all weekdays in April.

This list is passed on to another function, ‘generate\_results()’, which outputs three items:

* An excel workbook with NMBE values for each meter, considering predictions over all weekdays in April, for the 51 test cases (0 through 50 degradation days)
* A PDF with box plots of NMBE (over all meters) across the 51 testcases.
* A PDF with a scatter plot of median NMBE vs test case

1. **Export CSVs:**  export csv files with per hour prediction for each meter and each test case.

**Visualizations:**

**Shiny Dashboard:** interactive dashboard with two visualizations: NMBE boxplot for all test cases and a scatterplot of median NMBE against the test cases. Y-axis of the plot can be controlled by the user.

**Tableau Dashboard:** interactive dashboard showing a timeseries plot of energy predictions (of all three algorithms) and actual energy use over 24 hours. User control options include:

* Meter ID
* Number of degradation days
* Prediction date
* Algorithm

1. Net Mean Bias Error: sum(residuals)/n\*average(energy-kWh)). n refers to the number of datapoints [↑](#footnote-ref-1)