Predictability Analysis for Millennium Tower

09 February, 2021

## Predictability Analysis Purpose

The purpose of the model predictability analysis is to determine whether the building’s baseline energy use can be reliably determined over the duration of the project. Using empirical models, a building is considered predictable if:

* A model can be developed that meets certain goodness of fit and accuracy metrics as defined in this report, and
* The risk of non-routine events is limited or non-existent.

The purpose of this report is to document the findings of the model predictability analysis as a criterion for participation in the EEAS Program. This report documents the data received, the weather data station from which corresponding weather was selected, summary descriptions of the models developed, and a summary table of the key model goodness of fit and accuracy metrics for each model.

## Data Summary

|  |  |
| --- | --- |
|  | Utility Data Specification |
| eload | Electricity |
| Source | Seattle City & Light |
| Received Date | 02/02/2021 |
| Data Interval | 15 min |
| Start Date | 01/01/2019 00:15 |
| End Date | 12/31/2020 23:45 |
| # Missing Data Points | 0.0 |
| # Total Data Points | 69,965.0 |

|  |  |
| --- | --- |
|  | Temperature Data Specification |
| Weather Station | Boeing Airfield |
| Approx. Distance to Site | 5 mi |
| Data Download Date | 02/08/2021 |
| Data Interval | 60 min |
| Start Date | 01/01/2019 00:00 |
| End Date | 12/31/2020 23:00 |
| # Missing Data Points | 60.0 |
| # Total Data Points | 17,532.0 |
| TMY data available? | Yes |

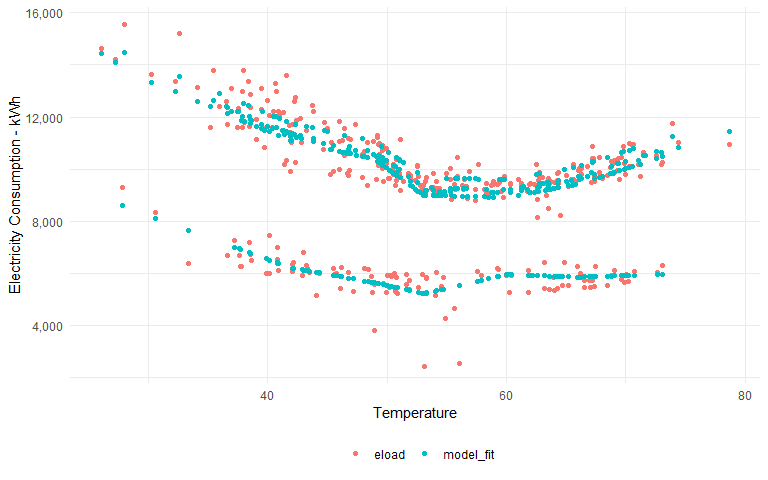
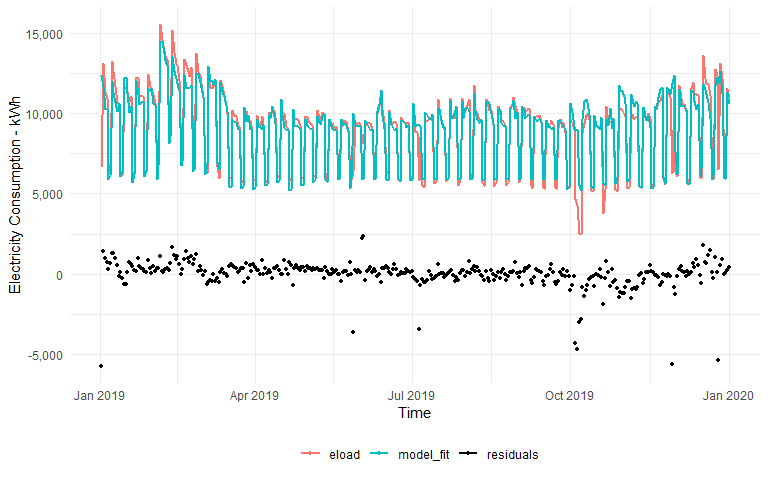
### Predictability Analysis Period

|  |  |
| --- | --- |
| Start Date | End Date |
| 01/01/2019 00:00 | 01/01/2020 00:00 |

## Training Period Models

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Interval | R^2 | Adjusted R^2 | CV(RMSE) % | NMBE % | Uncertainty for 10% Savings | Savings Frac for 50% Unc | CL |
| TOWT | Hourly | 0.89 | 0.89 | 14.68 | 0.000 | 0.15 | 0.03 | 90 |
| TOW | Hourly | 0.80 | 0.80 | 19.97 | 0.000 | 0.40 | 0.08 | 90 |
| Four\_P | Hourly | 0.08 | 0.08 | 42.56 | 0.000 | 0.57 | 0.11 | 90 |
| Three\_PH | Hourly | 0.03 | 0.03 | 43.57 | 0.000 | 0.59 | 0.12 | 90 |
| Three\_PC | Hourly | 0.03 | 0.03 | 43.61 | 0.000 | 0.61 | 0.12 | 90 |
| SLR | Hourly | 0.00 | 0.00 | 44.29 | 0.000 | 0.61 | 0.12 | 90 |
| TOWT | Daily | 0.86 | 0.85 | 9.95 | 0.000 | 0.12 | 0.02 | 90 |
| TOW | Daily | 0.70 | 0.69 | 14.51 | 0.000 | 0.42 | 0.08 | 90 |
| Four\_P | Daily | 0.15 | 0.14 | 24.32 | 0.000 | 0.29 | 0.06 | 90 |
| Three\_PH | Daily | 0.15 | 0.14 | 24.42 | 0.000 | 0.30 | 0.06 | 90 |
| HDD\_CDD | Daily | 0.13 | 0.12 | 24.70 | 0.000 | 0.30 | 0.06 | 90 |
| HDD | Daily | 0.09 | 0.08 | 25.18 | 0.000 | 0.30 | 0.06 | 90 |
| SLR | Daily | 0.07 | 0.06 | 25.43 | 0.000 | 0.31 | 0.06 | 90 |
| Three\_PC | Daily | 0.01 | 0.00 | 26.24 | 0.000 | 0.32 | 0.06 | 90 |
| CDD | Daily | 0.00 | -0.01 | 26.33 | 0.000 | 0.32 | 0.06 | 90 |
| Three\_PH | Monthly | 0.78 | 0.70 | 4.93 | 0.000 | 0.37 | 0.07 | 90 |
| Four\_P | Monthly | 0.80 | 0.69 | 5.04 | 0.000 | 0.38 | 0.08 | 90 |
| HDD\_CDD | Monthly | 0.70 | 0.59 | 5.79 | 0.000 | 0.43 | 0.09 | 90 |
| HDD | Monthly | 0.49 | 0.38 | 7.17 | 0.000 | 0.53 | 0.11 | 90 |
| SLR | Monthly | 0.46 | 0.34 | 7.36 | 0.000 | 0.54 | 0.11 | 90 |
| Three\_PC | Monthly | 0.23 | -0.06 | 9.28 | 0.000 | 0.69 | 0.14 | 90 |
| CDD | Monthly | 0.11 | -0.09 | 9.41 | 0.000 | 0.69 | 0.14 | 90 |

The model chosen for this project is TOWT.



## Assessment of NREs

***Add findings here***

## Goodness-of-Fit Metrics

**Coefficient of Variation of the root mean squared error, CV(RMSE)**

CV(RMSE) is a measure of random error between a model’s predictions and the actual data. Generally, we want to minimize this error as much as possible.

**Net Mean Bias Error (NMBE)**

NMBE is a measure of the difference between the model’s predictions of training period total energy use and the actual energy use. This error should be very low.

**Coefficient of Determination,**

The coefficient of determination describes how well the independent variables explain the variations in the dependent (energy) variable. Higher R2 means the independent variables have more explanatory power. This is an informative metric only, not a criterion, because while the energy use sometimes may not have high variation, an independent variable may adequately ‘explain’ the existing variation in the energy use, despite a low R2.

**Fractional Savings Uncertainty**

ASHRAE Guideline 14-2014 provided the following ‘fractional savings uncertainty’ formulas as a means to estimate the uncertainty of the savings estimated with this modeling approach. The formulas also enable the estimation of how well we will know savings based only on the baseline model’s goodness of fit, the number of points in the baseline and post-installation periods, the amount of savings, and the level of confidence at which we estimate the uncertainty. For daily or hourly models, they include a correction for autocorrelation. Using these formulas, we can estimate what the savings uncertainty, at 90% confidence, would be for a project that yields 10% savings, with a year of post-installation period monitoring, using a baseline model with its MSE or MSE’ value and a year of baseline data. We want the uncertainty to be low, but the minimum level of uncertainty cannot be greater than +/- 50% at the 90% confidence level. Note the percentage refers to the amount of savings, not to the baseline energy use.

Additional research by LBNL showed that ASHRAE’s formula underestimated uncertainty when used on hourly models, due to the high degree of autocorrelation in the data. This is why uncertainty in hourly models is not reported in the predictability report.

Savings Uncertainty, models with autocorrelation (hourly or daily):

where MSE’ is the mean squared error of the model.

Savings Uncertainty, models without autocorrelation (monthly):

Energy Savings Required for Uncertainty @ 90% Confidence Interval (10%) = 0.1\*E

Where:

* is the measured energy use in any time interval, in energy units
* is the model’s predicted energy use in any time interval, in energy units
* is the average energy use over all the time intervals, in energy units
* is the absolute precision of the savings estimate over m time periods, in energy units
* is the estimated energy savings over m time periods, in energy units
* is the number of data points in the training period
* is the number of parameters in the model
* is student’s t-statistic for the specified confidence level and n-p degrees of freedom
* is an equation depending on the analysis time interval:
  + for hourly interval data
  + for daily interval data
  + for monthly interval data
* M is the number of months in the reporting period
* is the number of data points in the model training period, corrected for autocorrelation
* is the number of data points in the proposed post-installation period
* is the expected savings, expressed as a fraction of training period energy use