# **1. Model Evaluation To characterize The Heat Load**

In this part we have explained the overall relationship between predictors and outcome variable. The best estimates for the coefficient to fit the model have been assessed. Moreover, a summary statistic table was presented in order to evaluate how the regression models fits the data. As regards to the graphical residual analysis, several plots were represented to check if our assumption were reasonable and whether the choice of model were appropriate or not.

## **1.1** **Feature Selection with Variance Thresholding**

Each of the four models were fitted to the ordinary least square regression (OLS). Thereafter a future selection using variance threshold was conducted. In this context, only the attributes that have significant effect on the model's output have been selected.

The selection procedure was based on p-value threshold in which the entry criterion was set to a value less than 0.05. To explain the variance in heat load outcome, the features were ranked by their individual ability, this means that each predictor that have a p-value greater 0.05 is excluded from the model. The models were refitted several times until no feature have a p-value greater than 0.05.

**Table 1. Regression models coefficients and P-value of significant variables (winter term)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Winter Term** | | | | | |
| **Workdays Model** | | | **Weekends Model** | | |
| **Predictor variables** | **Coefficients** | **P-value** | **Predictor variables** | **Coefficients** | **P-value** |
| POWER1\_lag1 | 0.1294 | 0.000 | POWER1\_lag1 | 0.1891 | 0.001 |
| POWER1\_lag2 | 0.0948 | 0.006 | POWER1\_lag2 | 0.0567 | 0.301 |
| POWER1\_lag3 | 0.0770 | 0.022 | POWER1\_lag3 | 0.1935 | 0.000 |
| POWER1\_lag12 | 0.1561 | 0.023 | Temperature\_lag0 | -0.5116 | 0.000 |
| Temperature\_lag0 | -1.0102 | 0.000 | Irradiation.flux\_lag0 | 0.0237 | 0.015 |
| Temperature\_lag1 | 0.9428 | 0.053 | Irradiation.flux\_lag1 | -0.0421 | 0.010 |
| Temperature\_lag2 | -0.3335 | 0.516 | Irradiation.flux\_lag2 | 0.0266 | 0.006 |
| Temperature\_lag3 | -0.9481 | 0.047 | Irradiation.flux\_lag11 | 0.0207 | 0.001 |
| Temperature\_lag4 | 0.7633 | 0.002 | Irradiation.flux\_lag12 | -0.0204 | 0.001 |
| Tuesday 1h | -2.1551 | 0.012 | Saturday 10h | 3.8600 | 0.000 |
| Friday 16h | 2.2787 | 0.014 | Sunday 10h | 2.4852 | 0.013 |
| Monday 8h | 2.5612 | 0.006 |  |  |  |

**Table 2. Regression models coefficients and P-value of significant variables (shoulder term)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Shoulder** | | | | | |
| **Workdays** | | | **Weekends** | | |
| **Predictor variables** | **Coefficients** | **P-value** | **Predictor variables** | **Coefficients** | **P-value** |
| POWER1\_lag1 | 0.2897 | 0.000 | POWER1\_lag1 | 0.2603 | 0.000 |
| POWER1\_lag2 | 0.2338 | 0.000 | POWER1\_lag2 | 0.2043 | 0.000 |
| Temperature\_lag0 | -0.6507 | 0.000 | POWER1\_lag3 | 0.1154 | 0.003 |
| Temperature\_lag1 | 0.2596 | 0.012 | Temperature\_lag0 | -0.3119 | 0.000 |
| Temperature\_lag24 | -0.0401 | 0.029 | Temperature\_lag24 | -0.0700 | 0.038 |
| Irradiation.flux\_lag1 | -0.0062 | 0.000 | Irradiation.flux\_lag4 | 0.0033 | 0.000 |
| Irradiation.flux\_lag2 | 0.0042 | 0.010 | Irradiation.flux\_lag24 | -0.0017 | 0.003 |
| Irradiation.flux\_lag3 | 7.809e-05 | 0.962 | Saturday 22h | 1.3799 | 0.027 |
| Irradiation.flux\_lag4 | -0.0020 | 0.222 | Sunday 10h | 1.2742 | 0.032 |
| Irradiation.flux\_lag5 | 0.0059 | 0.000 | Sunday 18h | 1.2186 | 0.038 |
| Irradiation.flux\_lag6 | -0.0029 | 0.005 | Sunday 22h | 1.2110 | 0.037 |
| Thursday 2h | -1.3630 | 0.025 |  |  |  |
| Friday 10h | -1.6718 | 0.003 |  |  |  |
| Friday 21h | 1.1456 | 0.044 |  |  |  |
| Monday 0h | -1.3003 | 0.032 |  |  |  |
| Monday 1h | -1.2724 | 0.037 |  |  |  |
| Monday 8h | 1.5540 | 0.011 |  |  |  |

**Table 3. Goodness of fit of the models with significant parameters using different evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Evaluation Metrics** | **Winter Models** | | **Shoulder Models** | |
| **Workdays** | **Weekends** | **Workdays** | **Weekends** |
| R-Squared | 0.845 | 0.763 | 0.883 | 0.866 |
| F-statistic | 342.7 | 76.03 | 707 | 359.7 |
| Prob (F-statistic) | 8.52e-296 | 5.68e-80 | 0.00 | 1.17e-258 |

To evaluate the accuracy of the models table 3. above shows the different evaluation metrics that are used to characterize the heat load. In this perspective the shoulder season models yield an excellent fit to the observed data. For workdays models 88% of the variability observed in the target variables is explained by the regression model.

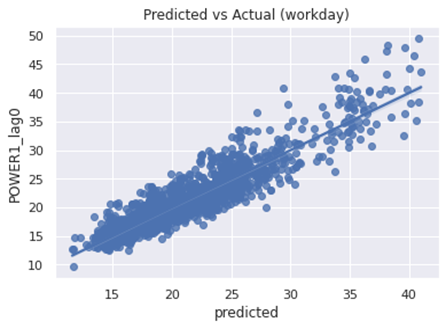
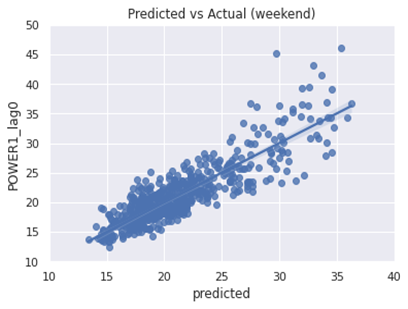
The minimum value of R-squared is shown by the winter model, essentially for weekends model where 76 % of the variability observed in the heat load variable is explained by the regression model.

The F-value and probability (F-statistics) depict the overall significance of the regression model and the accuracy of the null hypothesis respectively. In this context, the F- value of all the models were exceedingly large numbers compared to the F-critical value. (In our case study the F-critical value was equal to 1 for a significance level of 0.05 (α=0.05)). The models of the shoulder season displayed the higher F-statistic value particularly, in workdays model with a ratio of 707. For the prob (F statistic) we could see that all the values were exceedingly small numbers. For that reason, the null hypothesis of the models was rejected.

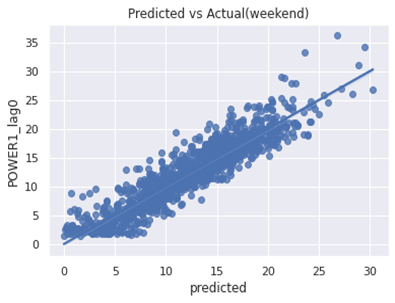
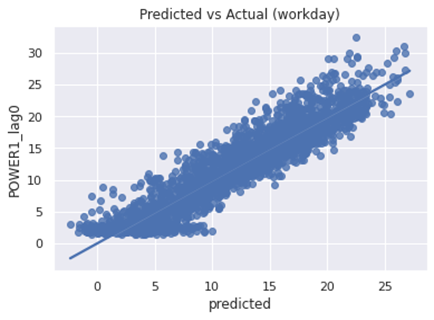
## **1.2** **Graphical Residual Analysis**

***Actual vs Predicted Data Graphs***

The charts below (fig 1) show the representation of predicted values of the heat load against the actual values. These types of graphs are used to see the level of accuracy of our models.



**(a) (b)**

****

**(c) (d)**

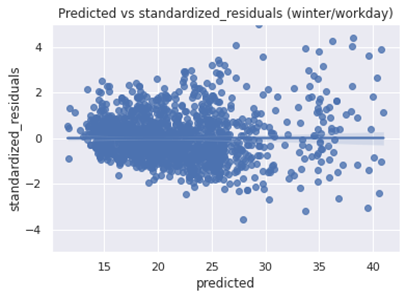
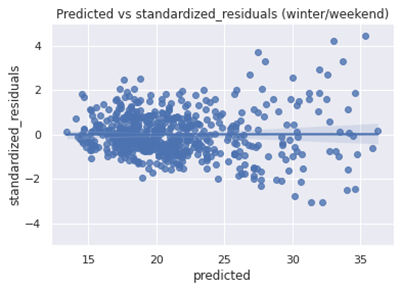
**Figure 1. Predicted data vs Observed heat load: (a) winter (workday), (b) winter (weekend), (c) shoulder (workday), (d) shoulder (weekend)**

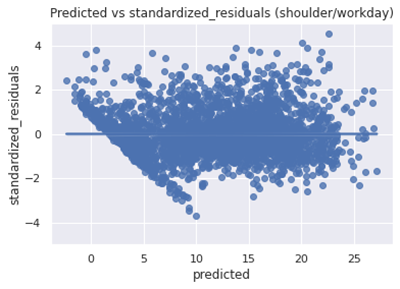
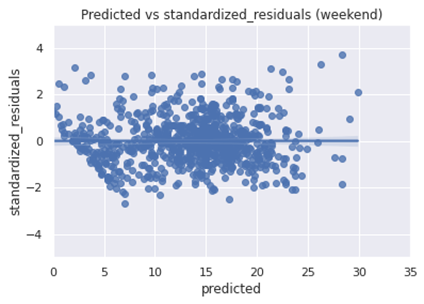
A strong correlation can be seen between model prediction and its actual results for shoulder term (both workdays and weekends) (c, d) compared to winter models (a, b). The roughly horizontal profile refers to the domestic hot water consumption (DHW) which has less to no dependence against weather data.

In workdays and weekends models of the winter term, presented in sub figures (a, b), The model has two subsections of performance: In the first subsection, where the actual values are between 10 to 25 kWh, the model does not seem bad. In the second subsection, the model values are between 30 to 40 kWh. Within this zone the goodness of fit of the model is weak and the points are dispersed.

***Residuals vs Predicted Data Graphs***

Plotting residuals versus predicted value of heat load (fig 2 below) enables to detect non-linearity or unequal error variance in the models. In general, the response of heat load should produce a distribution of points scattered randomly around 0.

 **(a) (b)**



**(c) (d)**

**Figure 2. Predicted heat load data vs Standardized residuals: (a) workday model (winter), (b) weekend model (winter), (c) workday model (shoulder), (d) weekend model (shoulder)**

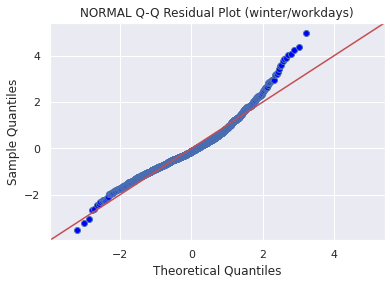
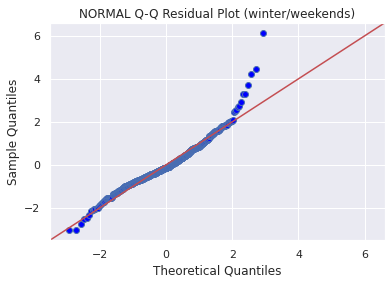
In winter term models we can see clear inequality in error variance and the models have a tunnel shaped structure. Indeed, in the first subsection, in which the heat load value is between 0 to 25 kWh, the two models show a symmetrically distribution of residuals which is one of the assumptions to have ideal plot. However, in the second subsection (zone between 30 to 40 kWh) we observe a change in the variance, where the fitted values are dispersed from the middle line. This type of behavior is known as heteroscedasticity. It happens essentially when the spread of the fitted values against the residuals is non-constant.

The models in the shoulder terms (sub figures c, d) display a symmetrically distribution of predicted heat load values against the residuals. Within this distribution the values are

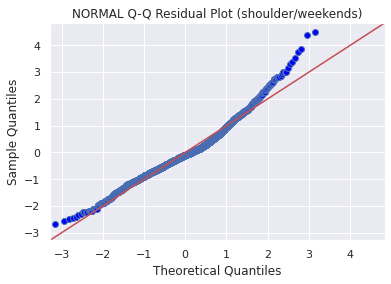
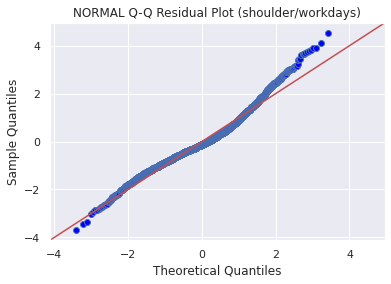
clustered toward the middle of the plot, nevertheless we see a structural pattern particularly in the first subsection within the zone of 0 to 10 kWh especially in the working model. This structural pattern suggest that the model has a room for improvement.

***Normal Q-Q Residual plots***

The purpose of the residual normal Quantile-Quantile (Q-Q) plot is to assess the data skewness and judge the normality distribution of the residual points. Fig 3. below displays the normal residual Q-Q plot for the four models.



**(a) (b)**



**(c) (d)**

**Figure 3. Normal Q-Q residual plots for winter and shoulder periods: (a) workday model (winter), (b) weekend model (winter), (c) workday model (shoulder), (d) weekend model (shoulder)**

In winter term the models (sub figures a, b) show a clear similarity in the Q-Q plot shape. Approximately, from the values (-2.5,2) the sample seems to grow at the same pace as the standard normal distribution. Consequently, their quantiles match this region and the points are aligned along the line. From the values (2,3) the residual sample grows slower than the theoretical quantile, therefore it reaches its highest quantile before the standard normal distribution. For this reason, we see a deviation from the line. This type of shape indicates that both two models have a room for improvement.

In shoulder season (sub figures c, d), approximately from the values (-3,2) the residuals sample grows slower than the standardized normal distribution, which it takes a longer time to increase. From (-2,1.5) The points are aligned along the line and the residual sample seems to grow at the same pace as the standard normal distribution. Lastly from the values (3,4) the sample grow faster than the standardized normal distribution, hence it reaches its highest quantile before the standard normal distribution. Typically, the S-shaped Q-Q plot indicate that the distribution has the correct median.