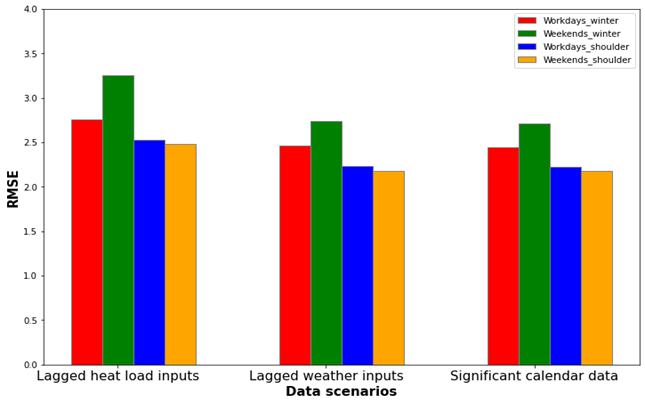
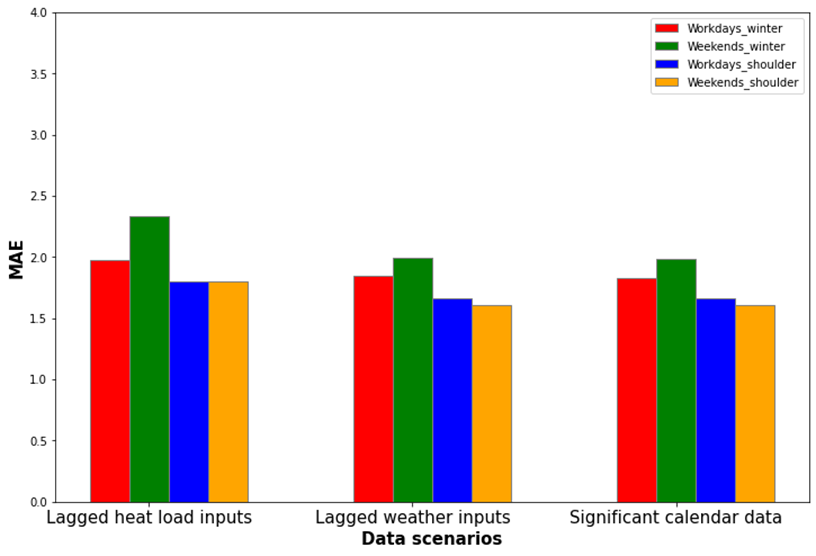
**1. Model Selection and Seasonal Performance Variation**

The model selection process has been performed in two main steps: First, among the various types of input data, the best data scenario that generate more accurate prediction was selected for each model as input parameters. Second, the selected models have been tested. Essentially, the seasonal performance variation of the heat load throughout the year were evaluated for each selected model.

## **1.1 Comparison of Models performance with Different Data Scenarios**



**Figure 1. Root mean square error of the four forecast models (workdays-winter, weekend-winter, workdays-shoulder and weekend-shoulder) for each data scenario**



**Figure 2. Mean absolute error of the four forecast models (workdays-winter, weekends-winter, workdays-shoulder and weekends-shoulder) for each data scenario**

From figure 1 and figure 2 above, we observe some variation between the magnitude of the mean absolute error (MAE) and the root mean square errors (RMSE). As a matter of fact, the RMSE is more sensitive to outliers and penalizes the larger error that's why the value of root mean square errors are higher than the mean absolute errors.

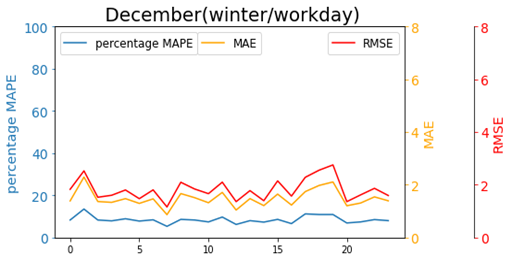
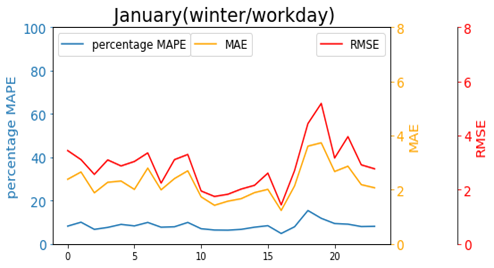
Models with only lagged heat load variables stands out by performing significantly worse than models with other data scenarios (lagged weather variables and significant calendar data).

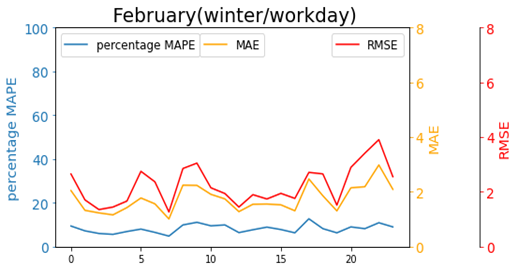
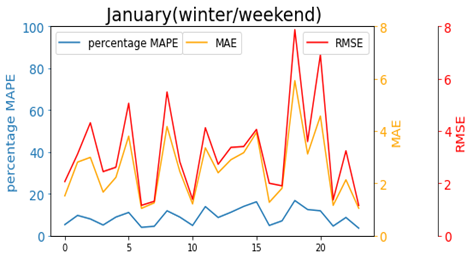
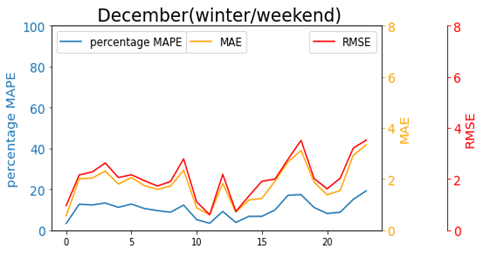
In this perspective, winter models have the highest values of RMSE and MAE compared to shoulder models in all the three data scenarios. (Particularly the weekend model in winter terms with an RMSE value of 3.254 kWh and an MAE value of 2.334 kWh).

Adding lagged weather variables of temperature and irradiation flux drastically increase the performance of all the four models which is explained by the decreasing value of the mean absolute error (MAE) and the root mean square error (RMSE).

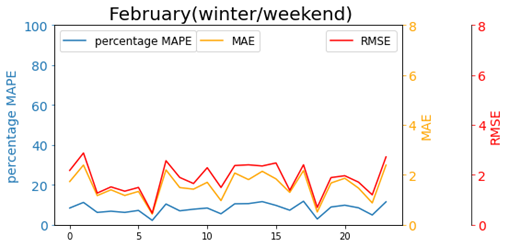
One can see in both figures that adding the significant calendar data slightly improve the performance of the models. **The focus in the rest of this work will be on the four models of the 3rd data scenario using significant calendar data.**

## **1.2 Seasonal Performance Variation**

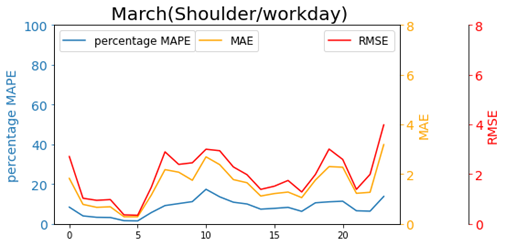
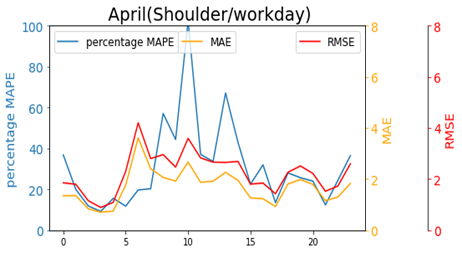
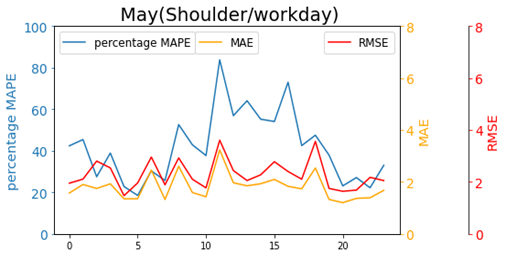
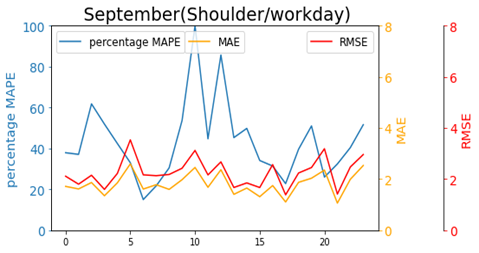


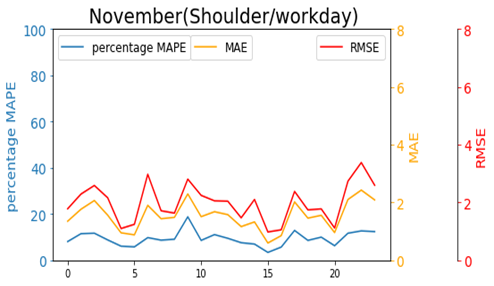
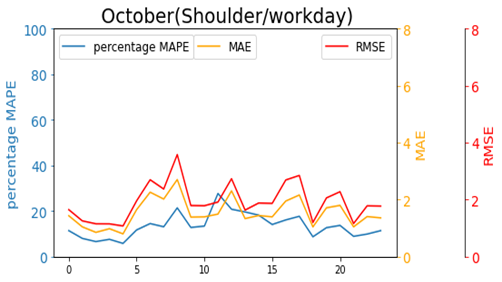


**(A)**

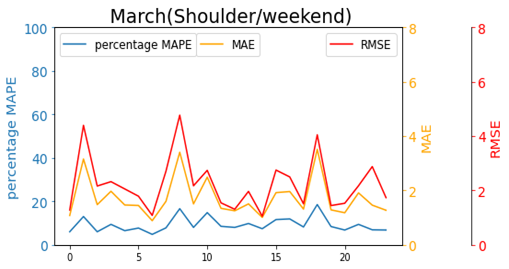
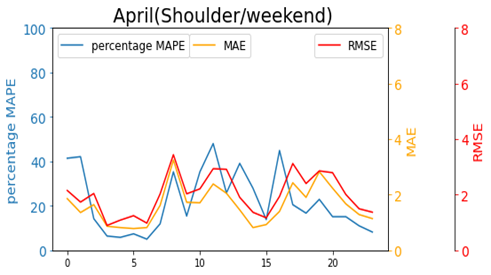


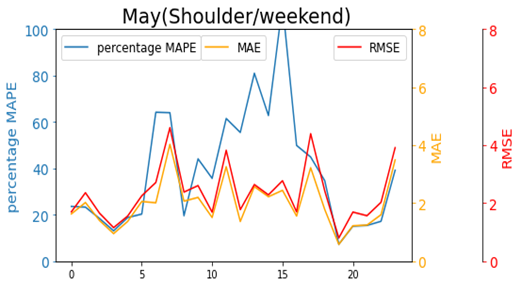
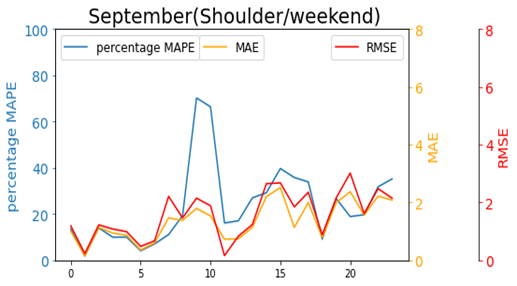
**(B)**

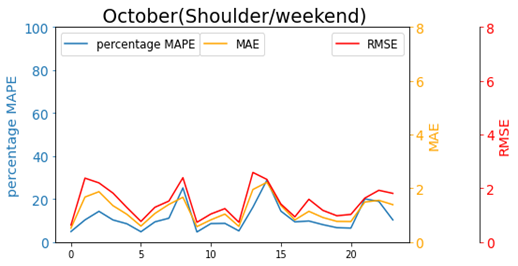
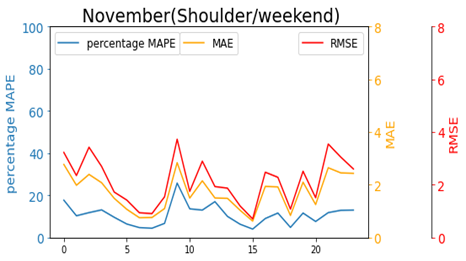
****



**(C)**

****

****

****

**(D)**

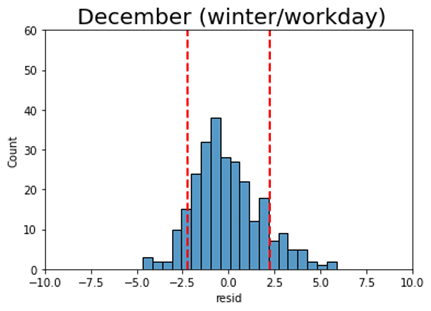
**Figure 3. Performance of the four models in each month, where the forecast error varies with time of the days, using the significant calendar data: (A) winter / workdays model, (B) winter / weekend model, (C) shoulder /workdays model, (D) shoulder /weekend model**

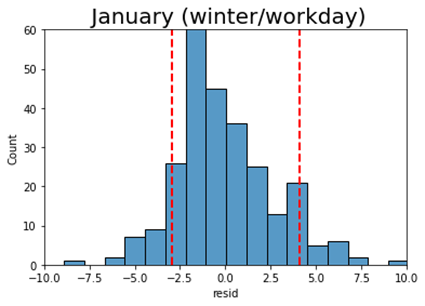
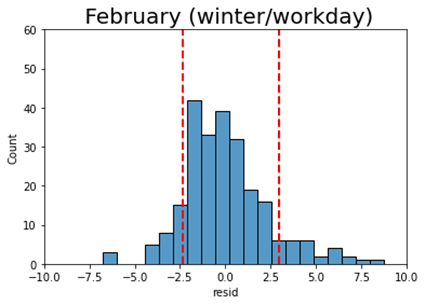
As shown in figure 3. Above the heat load varies over the year in magnitude and variation. Three different evaluation metrics are shown: on the left axis the mean absolute percentage error (MAPE), on the right axis the mean absolute error (MAE) and root mean square error (RMSE). The hours of the day are represented by the horizontal axis.

We can see that the RMSE and MAE appear to be large in winter term compared to shoulder months. In this context the evaluation metrics are large in the evening hours, this could be explained by the large heat load with large variance during the winter months. The RMSE and MAE can be above 4 kWh and 3 kWh in some hours especially in January whereas in shoulder term it can be below 2 kWh especially in April.

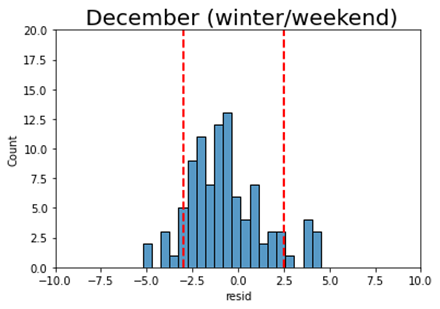
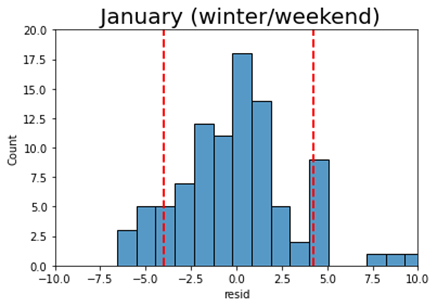
The MAPE behaves in opposite way compared to RMSE and MAE metrics, generally the aforementioned metric is smaller in winter months with a value between 4 to 20 % and larger in shoulder months with values between 20 to 80 %. These large values of MAPE indicate the inaccuracy of forecasting in some hours of shoulder months especially for April, May and September where we observe a peak of MAPE particularly between 10:0 00 to 15:00.

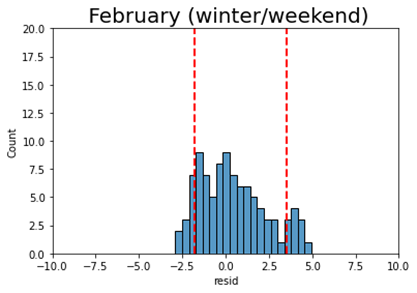
As a whole we didn't spot any clear pattern of error change during the day. Moreover, the performance of the four models doesn't seem to be bad in early morning particularly between 3:00 to 4:00.

****

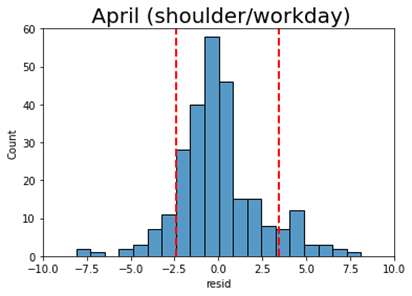
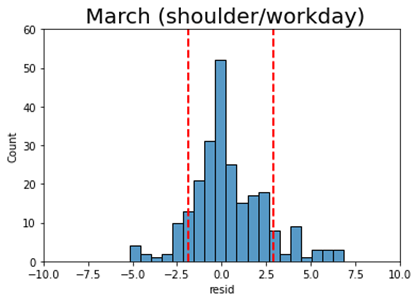
****

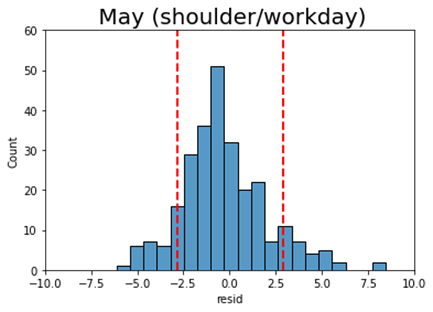
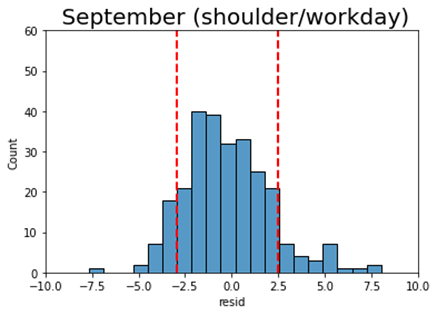
**(A)**

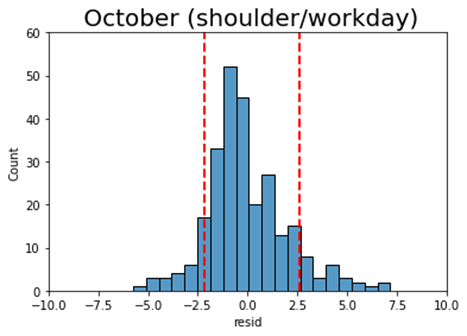
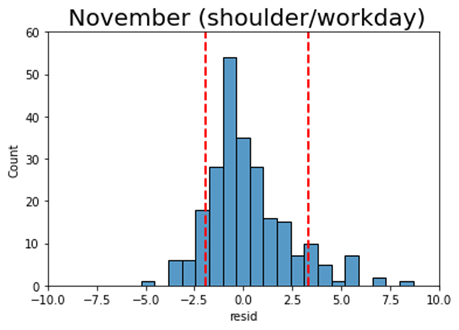




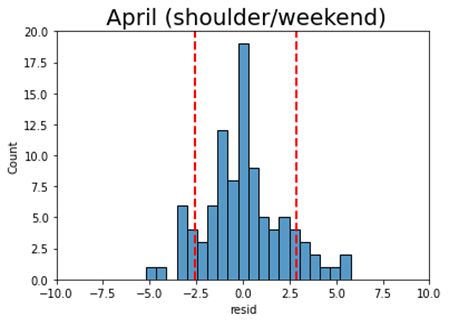
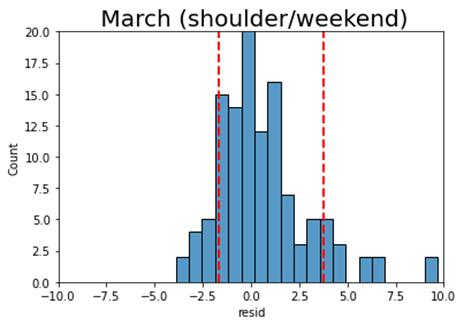
**(B)**

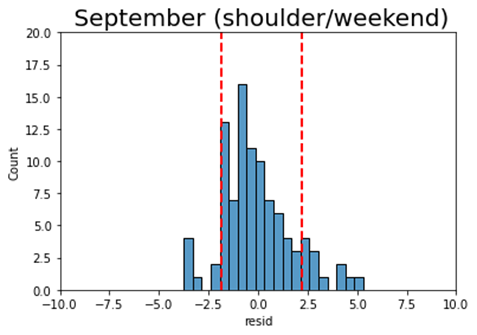
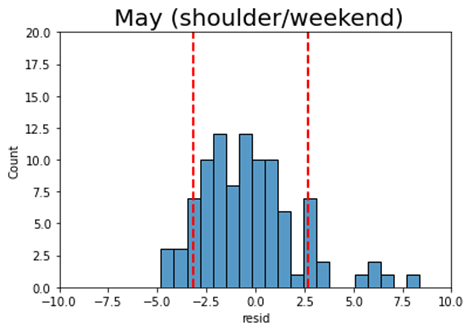
****

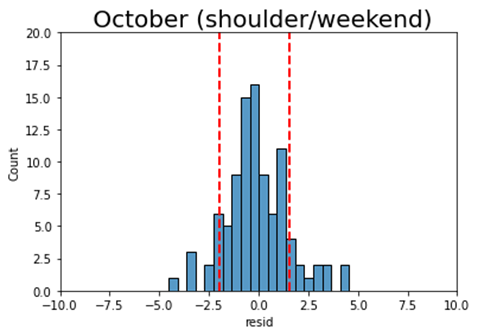
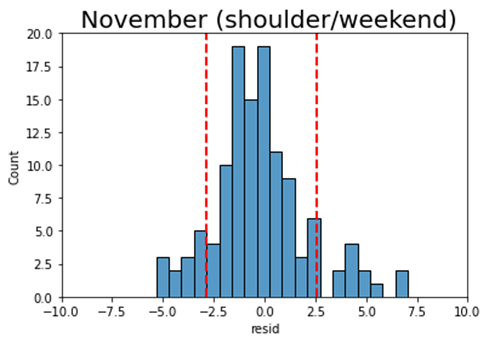




**(c)**

****





**(D)**

**Figure 4. Histogram of the forecast error of the four models using significant calendar data. The forecast error distribution is shown for each month in the year along with 10% and 90% quantiles. The forecast is too high when having a positive error, and too low when having negative error: (A) winter /workdays model, (B) winter / weekend model, (C) shoulder /workdays model, (D) shoulder / weekend model**

The aggregated error metrics such as RMSE, MAE and MAPE do not tell us the full story regarding the forecast variation over the year. As a matter of fact, identifying maximum error can be helpful in the heat load production planning and also can ensure the evaluation risk to cover the heating load day by day.

Figure 4 above represents histogram for the hourly errors for each month. The 10% and 90% quantiles are specified. In this context the width of the error’s distribution varies from month to month.

In the shoulder season the forecast errors are more or less confined particularly in October of weekend model. While in the winter term the distribution errors widen specifically in January month of the weekend model. In addition, the distribution of errors is not completely symmetric around 0. In March (workday model), for example, the distribution is shifted to the positive side. However, in May (weekend model) the distribution is shifted to the negative side.

**Table 1. Summary of the hourly forecast error for each month of the four models using significant calendar data. The best month performance is indicated in green, the worst is indicated in red. The quantiles are evaluated in pairs, where the month that have the widest quantile interval is considered the worst.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **RMSE** | **ME** | **Error Quantiles (kWh)** | | | |
|  |  | **(kWh)** | **(kWh)** | **10%** | **90%** | **1%** | **99%** |
| **Winter/ Workdays model** | December | 1.865 | -0.064 | -2.238 | 2.283 | -4.079 | 4.694 |
| January | 3.000 | 0.055 | -2.969 | 4.092 | -5.789 | 9.702 |
| February | 2.339 | 0.009 | -2.321 | 2.966 | -5.240 | 6.595 |
| **Winter/ Weekends model** | December | 2.188 | -0.521 | -3.008 | 2.491 | -4.857 | 4.430 |
| January | 3.665 | 0.228 | -3.992 | 4.220 | -6.134 | 13.509 |
| February | 1.959 | 0.434 | -1.771 | 3.550 | -2.601 | 4.424 |
| **Shoulder/ Workdays model** | March | 2.142 | 0.384 | -1.893 | 2.935 | -4.979 | 6.408 |
| April | 2.373 | 0.005 | -2.411 | 3.478 | -5.887 | 6.619 |
| May | 2.359 | -0.284 | -2.799 | 2.936 | -5.201 | 5.909 |
| September | 2.322 | -0.226 | -2.963 | 2.486 | -4.524 | 6.371 |
| October | 2.029 | -0.048 | -2.175 | 2.610 | -4.605 | 5.830 |
| November | 2.099 | 0.225 | -1.954 | 3.319 | -3.564 | 6.237 |
| **Shoulder/ Weekends model** | March | 2.449 | 0.594 | -1.694 | 3.749 | -3.371 | 8.713 |
| April | 2.127 | 0.028 | -2.586 | 2.895 | -4.324 | 5.618 |
| May | 2.542 | -0.333 | -3.167 | 2.718 | -4.719 | 6.691 |
| September | 1.766 | -0.086 | -1.823 | 2.229 | -3.627 | 4.668 |
| October | 1.589 | -0.104 | -1.953 | 1.553 | -3.469 | 4.537 |
| November | 2.311 | -0.204 | -2.864 | 2.595 | -4.944 | 6.557 |

In table 1. above a summary of the error distribution is represented. The 99% and 1% quantiles indicate the maximum errors. In winter term the best month of workdays model is December with 98% falling between -4.07 kWh and 4.69 kWh, while January is the worst month where 99% of errors are below 9.70 kWh and 1% of data are less or equal -5.78 kWh.

The best month of weekend model is February with 98% of the errors falling between -2.60 kWh and 4.42 kWh. while January, also here, is the worst month where 99% of errors are below 13.5 kWh and 1% are less or equal -6.13 kWh.

In shoulder season the best month of workday model is October with 98% of the errors falling between -4.6 kWh and 5.8 kWh. In contrast April is considered as the worst month where 99% of errors are below 6.6 kWh and 1% are less than -5.8 kWh.

As the workday model, October represents the best month for the weekend model with 98% of errors falling between -3.4 kWh and 4.5 kWh and the worst model is identified in March where 99% of errors are less or equal 8.7 kWh and 1% are less than -3.3 kWh.

Additionally, the forecast errors are biased differently in each month for each model. In our case and according to the mean error metric (ME), we did not observe any clear bias in the four models.

When evaluating the performance of models using the RMSE metric we can say that January has the largest value in winter term for workday and weekend model, with being the hardest month to forecast. However, December and February show the smallest value of RMSE in workday and weekend models respectively, with being the easiest months to forecast.

In shoulder period October has the smallest value for both workday and weekend model with being the easiest month to forecast. Whereas May shows the largest value of RMSE for both workday and weekend models, with being the hardest month to forecast.