We can prove significant outperformance using machine learning

~28 %

Improved Prediction[1]

(compared to presistence model)

~770 £/h

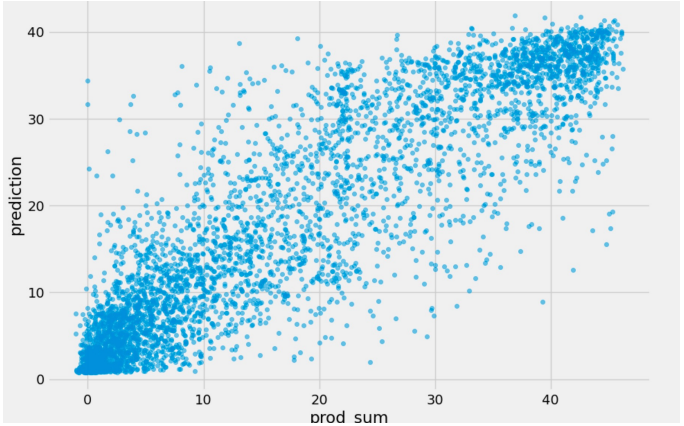
Possible Savings[2]

(compared to presistence model)

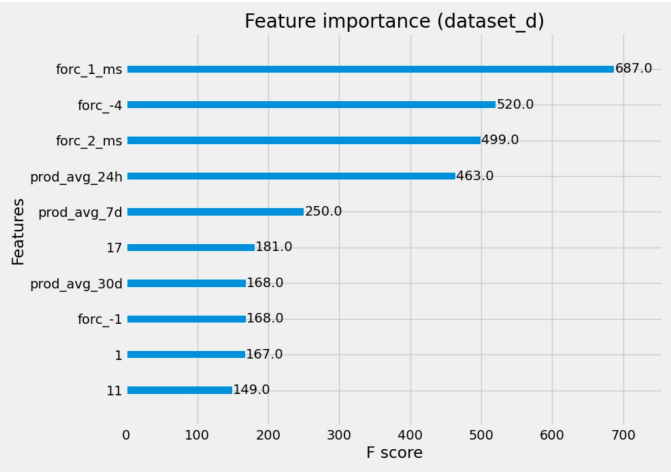
Out-of-sample[3]

Method	RMSE	Adj.R2	RMSE Improv.[1]
Persistence	9.2132	0.6061	-
OLS	7.7578	0.7207	16%
XGB[4] – Dataset A	6.983	0.7727	24%
XGB – Dataset B	6.9054	0.7794	25%
XGB – Dataset C	6.9662	0.7722	24%
XGB – Dataset D	**6.6513**	**0.7918**	**28%**

Predicted vs. Actual



Feature Importance



Data Cleaning

DATASET A
- 20 Power
- 2 Wind Forecast
- Date & Time

DATASET B
- ***Power cleaned***
- Wind Forecast
- Date & Time

DATASET C
- Power cleaned
- Wind Forecast
- ***Avg, Min, Max per Month***

DATASET D
- Power cleaned
- Wind Forecast
- Avg, Min, Max per Month
- ***Lagged Power, Forecasts***

[1] with XGB and dataset D compared to persistence model with dataset A
[2] Assumption that 1 MW/h can be sold for approx. £180.00/MWh with avg. production of 15.25MW/h. Source: https://marketwatch.zenergi.co.uk/price/22-02-2022/
[3] Each dataset was split into in-sample (75%) and out.of-sample (25%)
[4] Extreme Gradient Boosting (XGB)