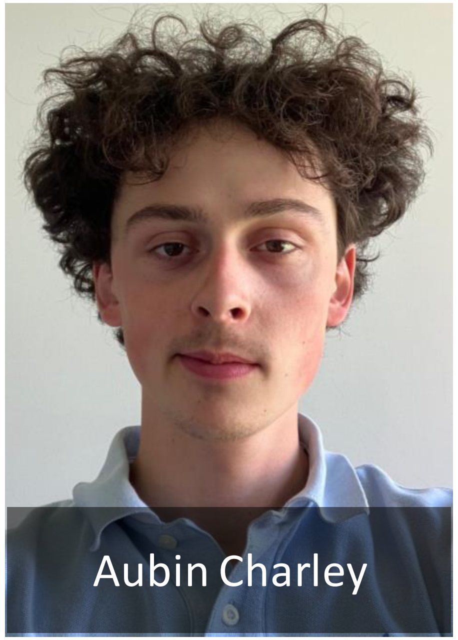
EDF & IOT
Forecasting electricity consumption using AI

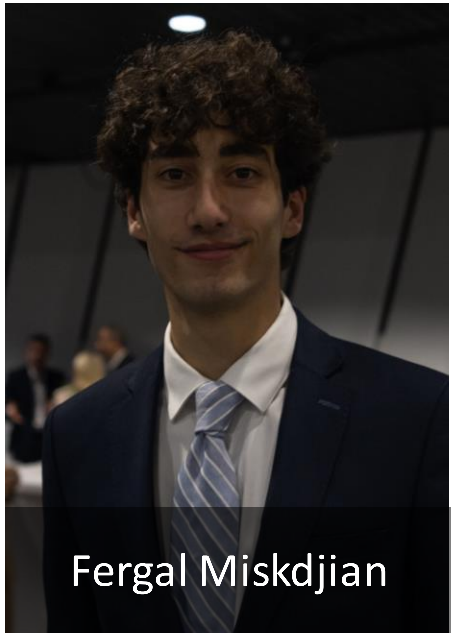
The Forecasters


Jilani Rim


Tarik Ouajdou


Wandrille Joudinaud


Aubin Charley


Fergal Miskdjian

The Team

Summary

Context

Ecological transition => efficience and sobriety

Methodology

Optimization of our prediction in terms of RMSE adding one by one new parameters

Techniques

Data analysis

Machine learning

Electricity networks in the ecological transition

Goal : Forecasting consumption to monitor production

Forecasts adjusted in real time with meteorological data

Used to decide

- Importations
- Exportations
- Pumping
- Adjust production

Essential to:

- Equilibrium
- Safety

Data collection

Sources of the data

An open platform of the French government for the calendar features

A weather station network for the climatic features

A selection of index meter reader eventually corrected

Key variables collected for the analysis :

Date and hour

Temperature

Nebulosity

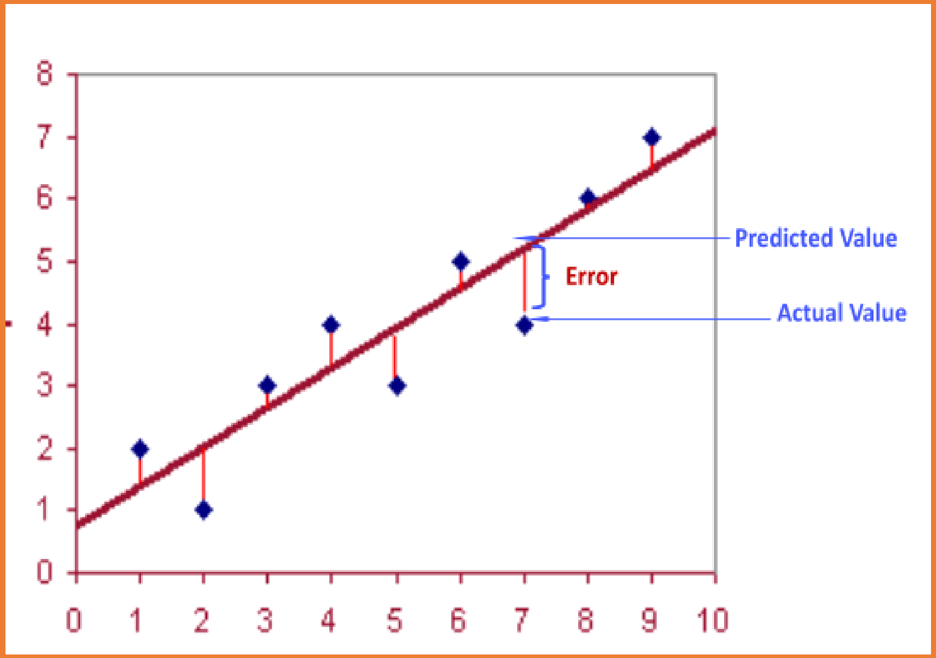
Wind

Correlation between the different parameters and not controllable balance energy

High score => high correlation

	Features	Score
1	year	221.901258
3	toy	10267.347391
6	temperature	16514.676243
4	week_number	22457.065407
8	wind	24145.801151
12	period_holiday	28706.811152
11	day_type_week_jf	40819.204994
10	day_type_jf	44522.856336
5	period_hour_changed	46546.877915
14	period_summer	48467.300457
13	period_christmas	51822.899559
9	day_type_week	71305.658252
0	month	113373.028242
7	nebulosity	211466.236588
15	nebulosity_by_solar_power_weights	219731.254405
2	tod	425438.774038

Model development : training and testing



Evaluation metrics : RMSE

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

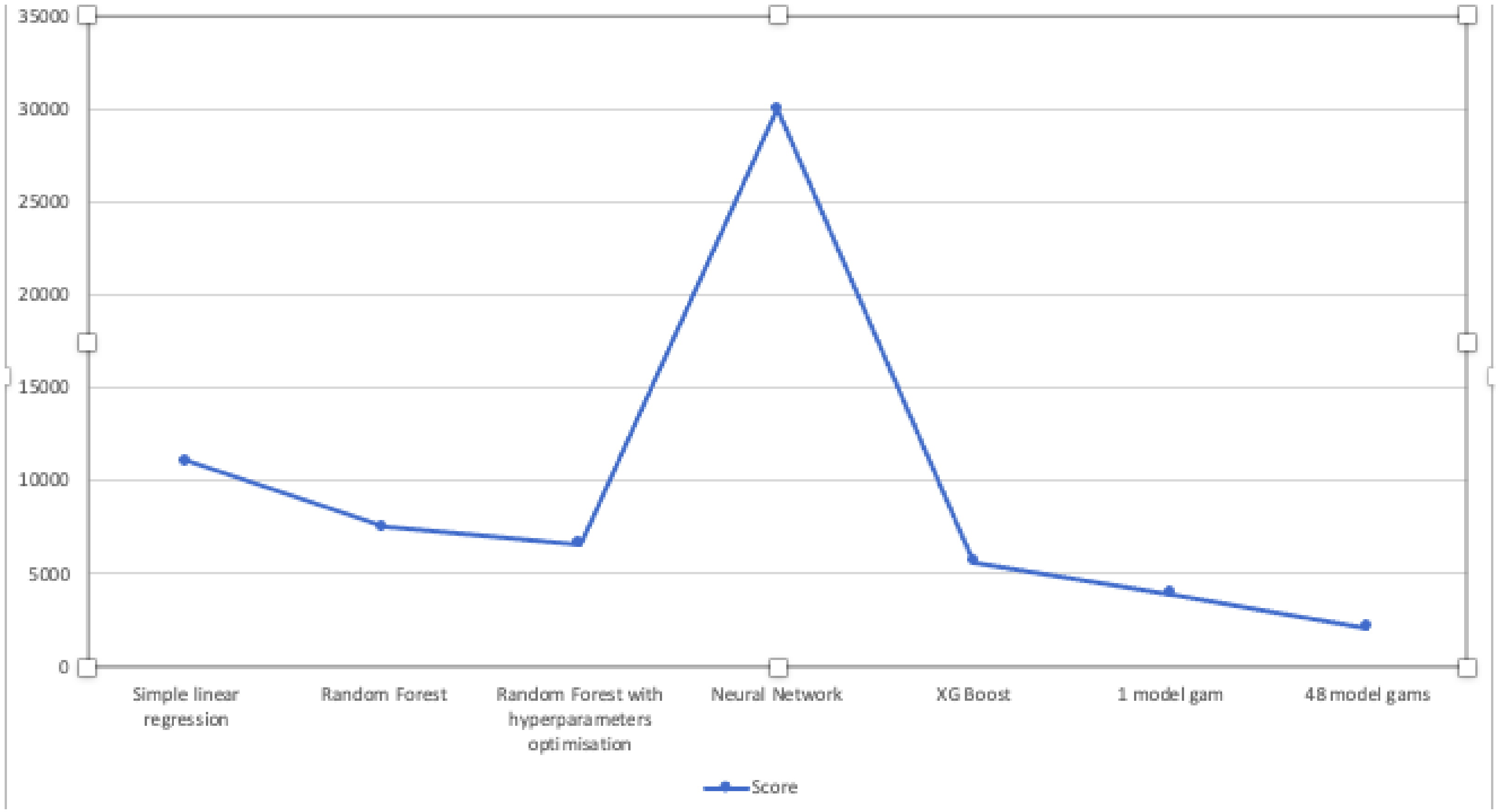
$\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are predicted values

$y_1, y_2, \ldots, y_n$ are observed values

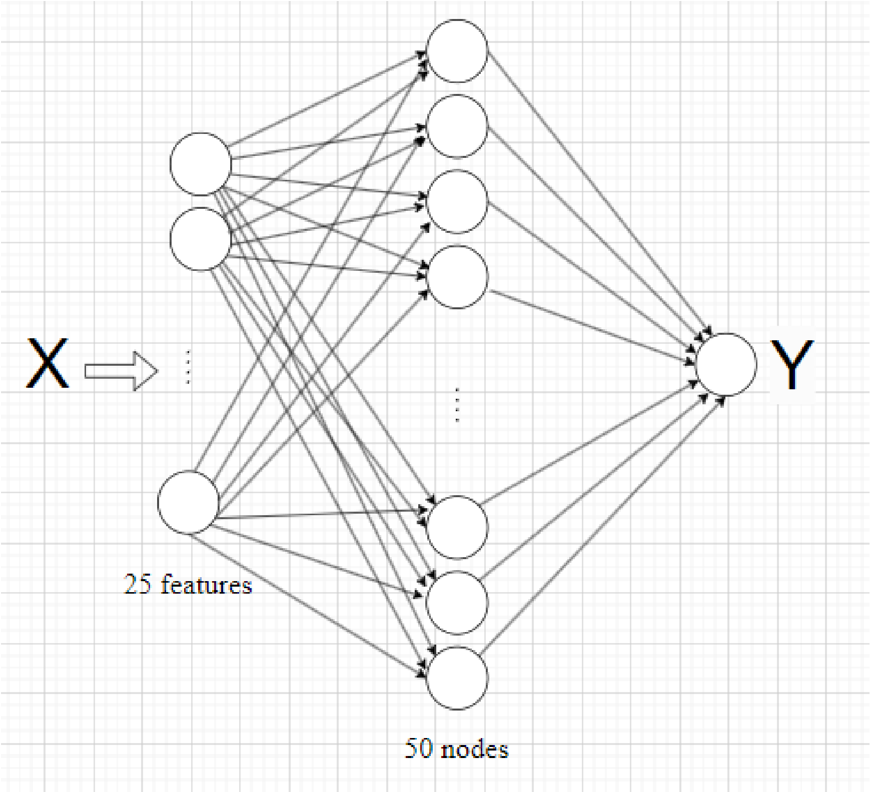
n is the number of observations

$$\mathrm{RMSE}_{\mathrm{combined}} \approx \sqrt{\mathrm{RMSE}_1^2 + \mathrm{RMSE}_2^2 + \mathrm{RMSE}_3^2}$$

Our different trials



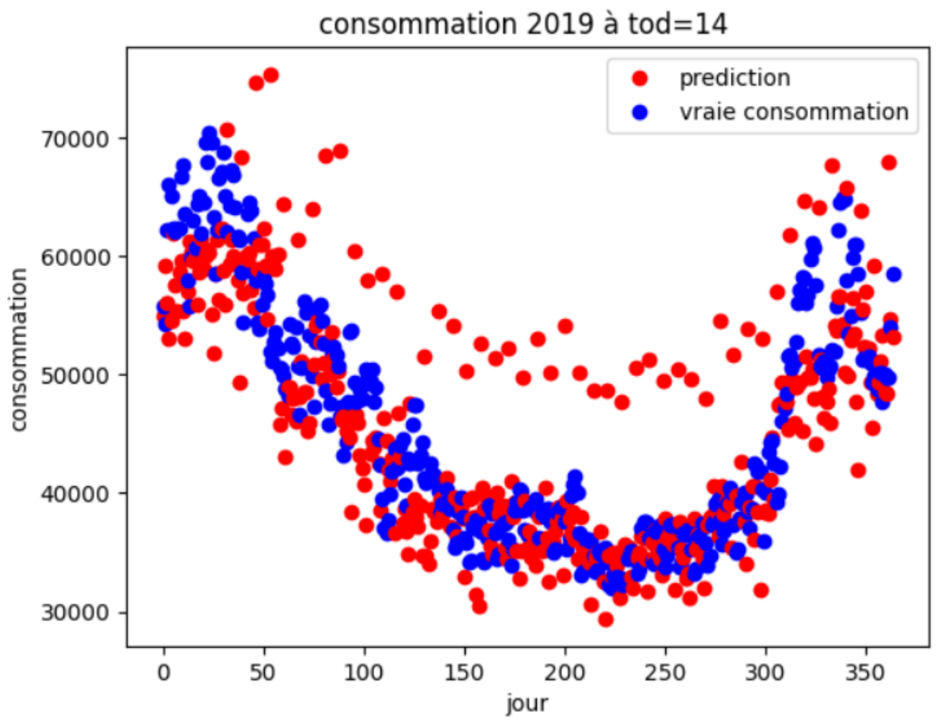
Methods and approaches : Neural Network

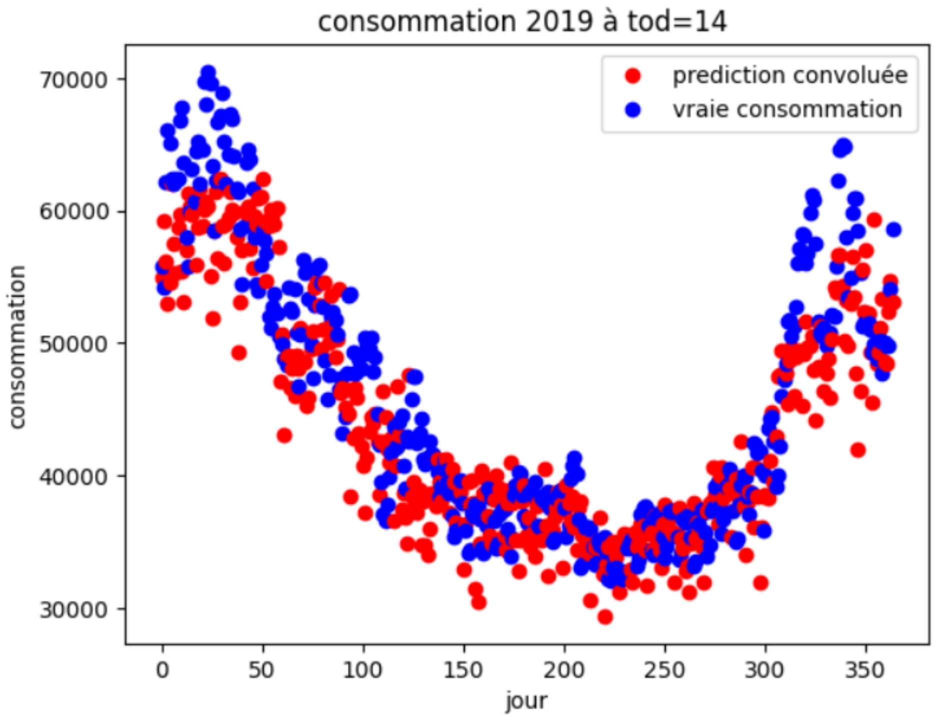


PyTorch

- Activation functions
- Choice of features: following previous analysis
- Random choice of hyperparameters
- Normalised data
- Dropout
- One-Hot encoding

Results of Neural Network simulations



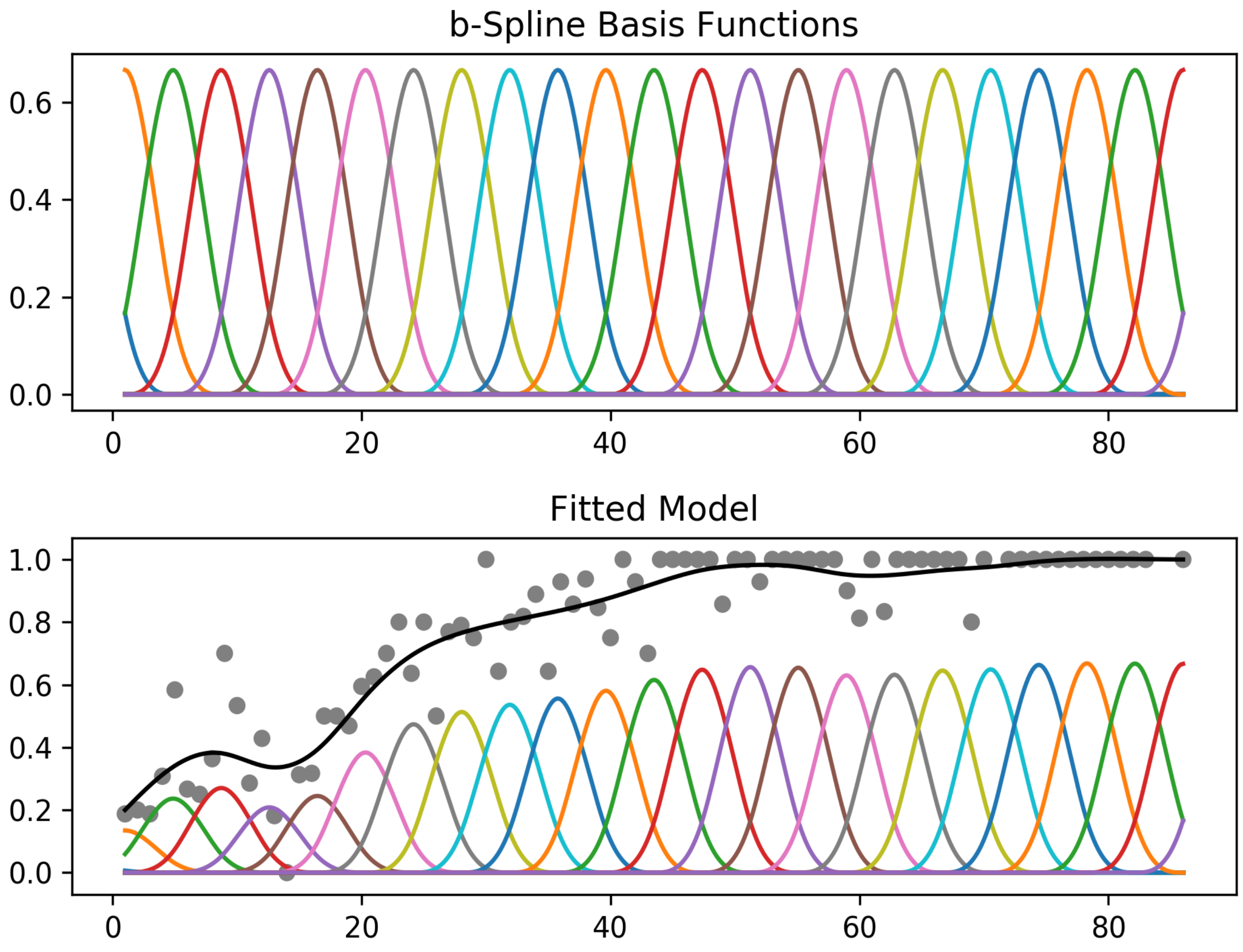


$$y_{conv}(i+1) = \begin{cases} \alpha y(i+1) + (1-\alpha) y_{conv}(i) & \text{si } |y_{conv}(i) - y(i+1)| \geq seuil \\ y(i+1) & \text{sinon.} \end{cases}$$

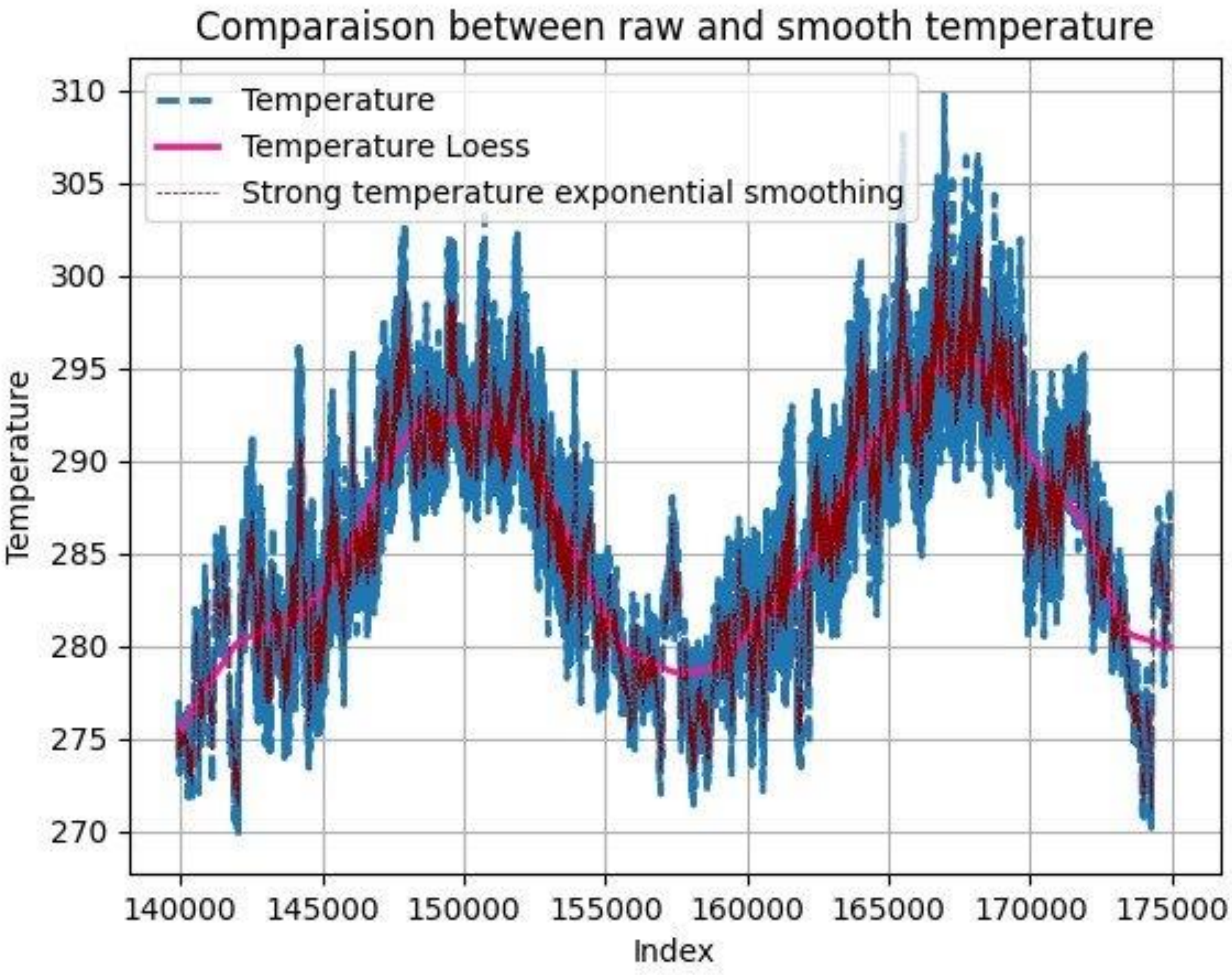
Methods and approaches : Linear regressions

- The linear regression : We developed 48 linear regression models to predict each half-hour interval of the day separately

- Other data considered :
    - Covid
    - Sobriety
    - Smoothed temperatures
    - Shifted temperatures



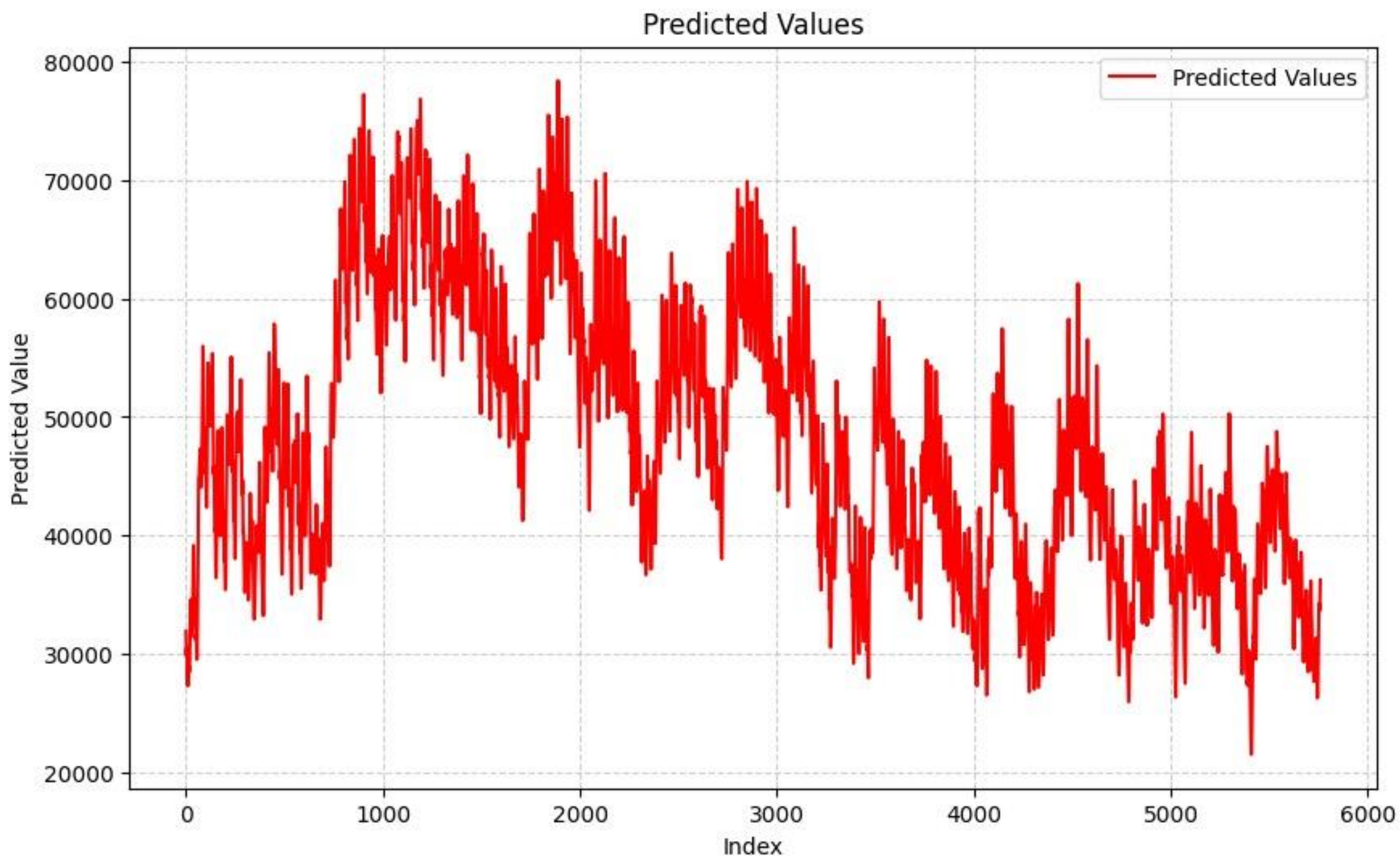
Temperatures Smoothing

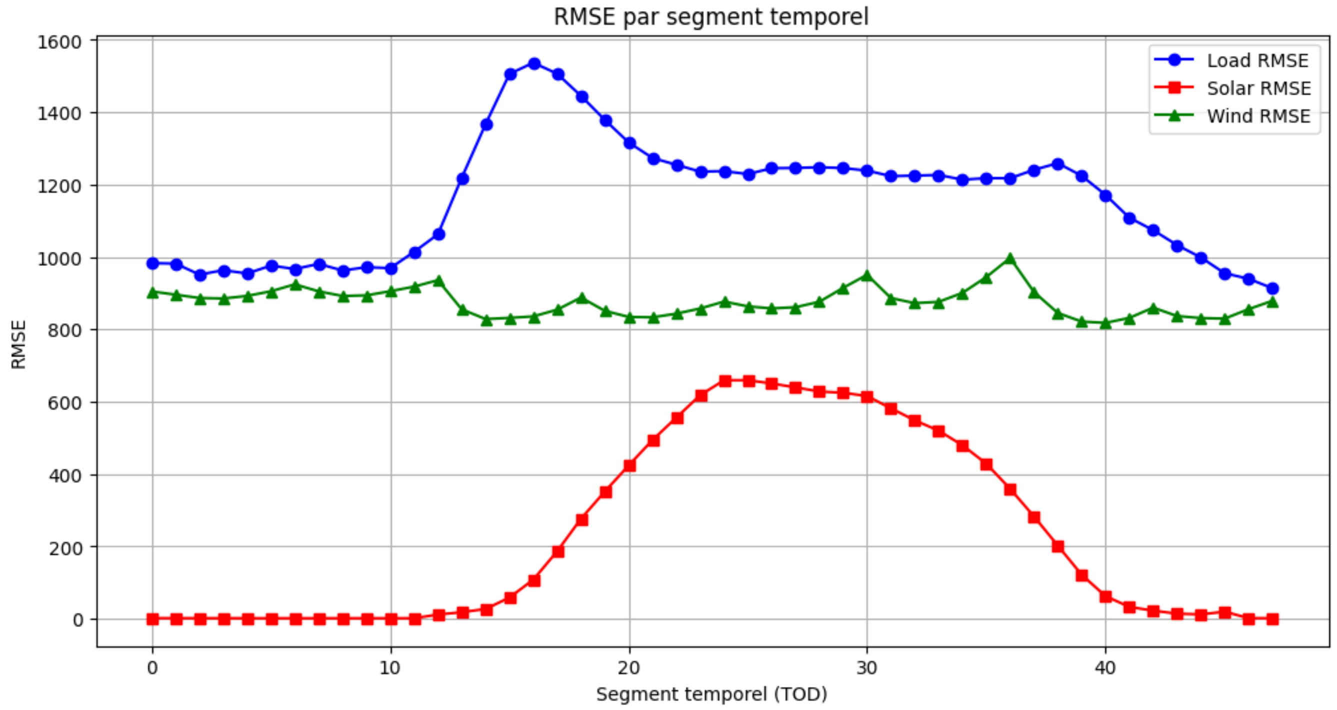


LOESS: LOcally Estimated Scatterplot Smoothing

Results

- Predicted values for 2023





Results of the simulations

Challenges

Data quality

Model improvements

Conclusive Insights

The linear regression models provided reasonably accurate forecasts

- Potential improvements and future work :

-different linear regressions for each day of the week

-real time model adjustment

-more advanced machine learning models

- Imagining an IoT system to have perfect energetic balance in the electrical network

Questions and Answers