Predict energy consumption and generation

Baigalmaa BAATAR

Objective

Build a machine learning model that predicts the production and energy consumption of a solar installation.s

Project idea

Future of energy:

The world is in a transition to new kinds of energy:

Blockchain enabled software platform for trading renewable energy and environmental commodities.

“Energy independence is new independence.”

The peer-to-peer nature of blockchain could provide a particularly useful answer to the existing problems we see within energy networks around the world.(Power Ledger company)

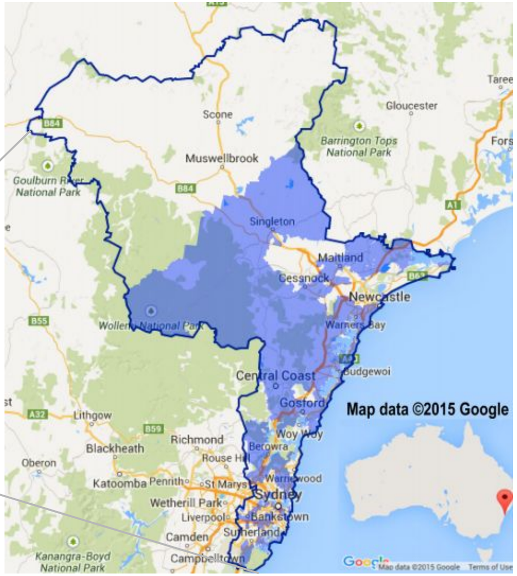
Image source

Contents

Collecting data/Web scraping

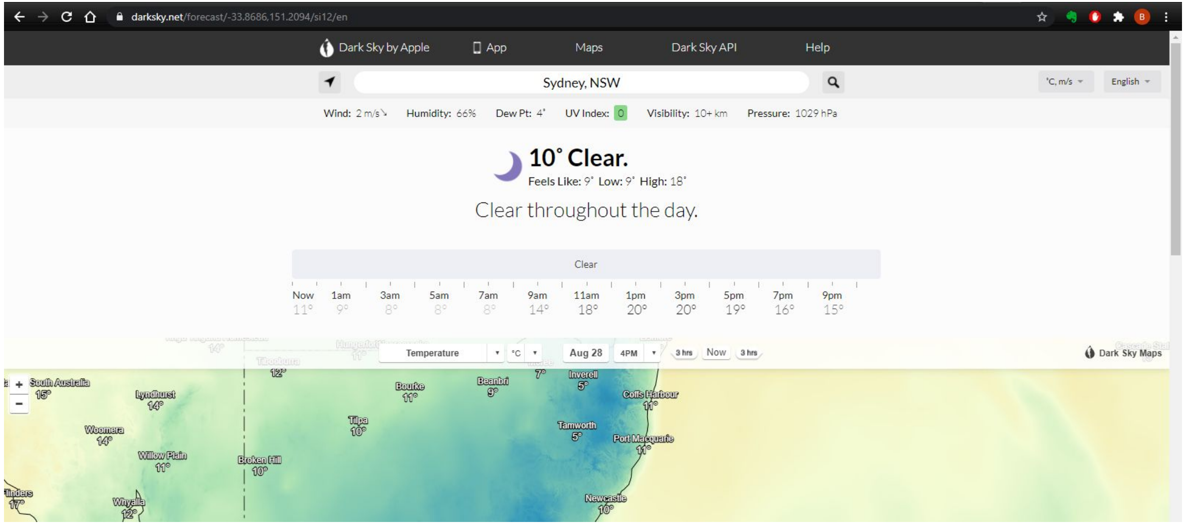
Data

The data for this project comes from Ausgrid's API and covers 300 homes with solar rooftop systems.



Weather information for energy generation

I used weather information from DarkSky.net. Web scraping to obtain 3 years of weather data, temperature, UV index, snow, in total 14 features.



Data overview

Location	Australia, East coast
Date	2010-2013
Number of observations	300 household (52602 observations)
Repeat	Every 30 mins
Target	Energy Consumption(GC)

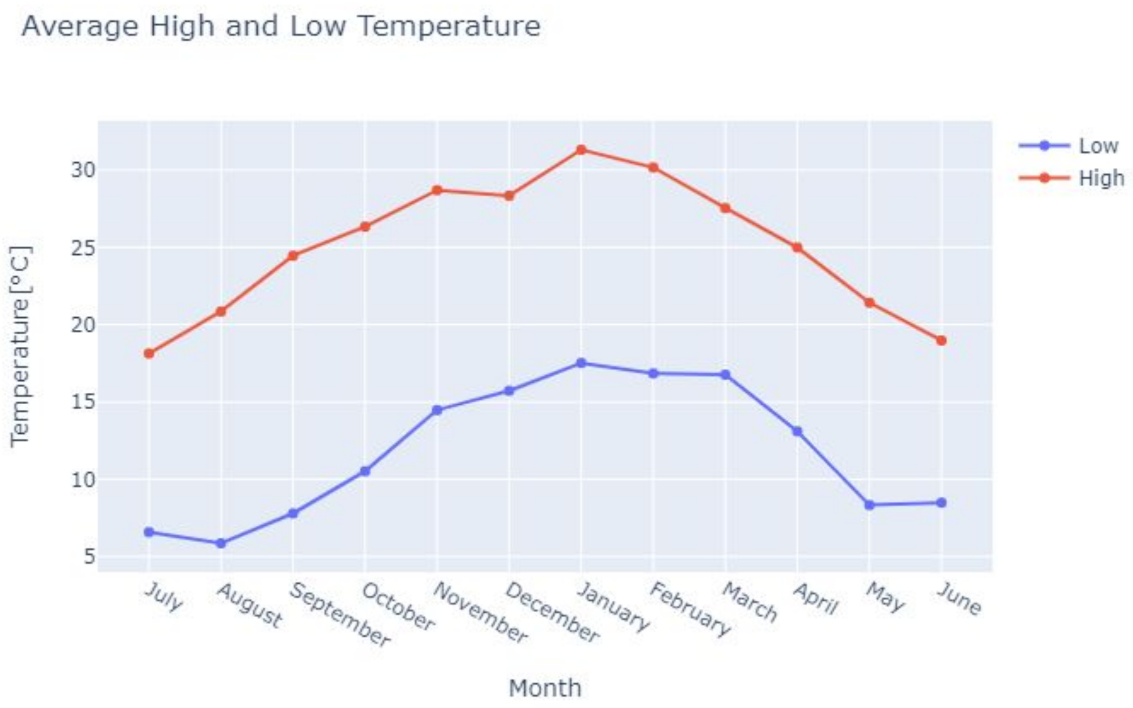
Data for energy consumption

Location	Australia, East coast
Date	2010-2013
Number of observations	300 household (52602 observations)
Repeat	Every 30 mins
Features	Temperature, Humidity,UV Index, Cloud Cover, Rain chance, Snow chance, Dew point, Pressure, Wind speed, Wind Gust, Wind Bearing, Visibility
Target	Energy Generation(GG)

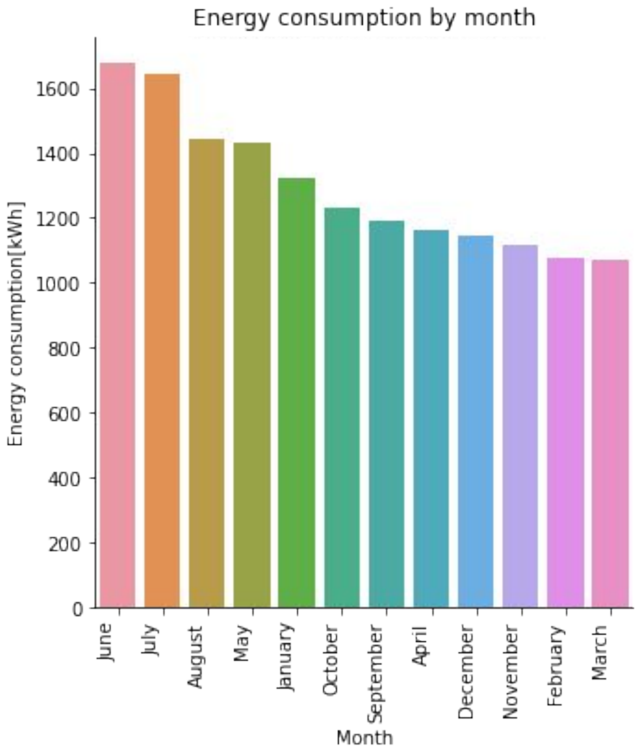
Data for energy generation

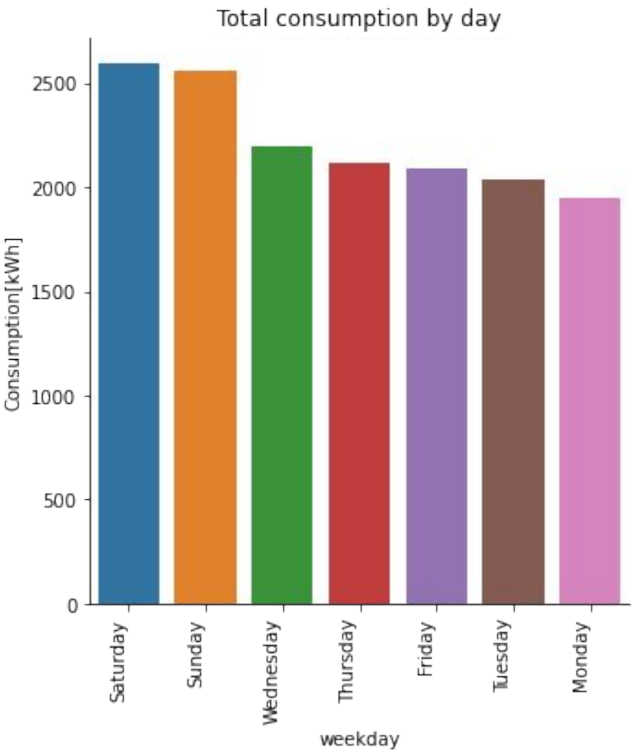
EDA(Exploratory Data Analysis)

Average High and Low Temperature

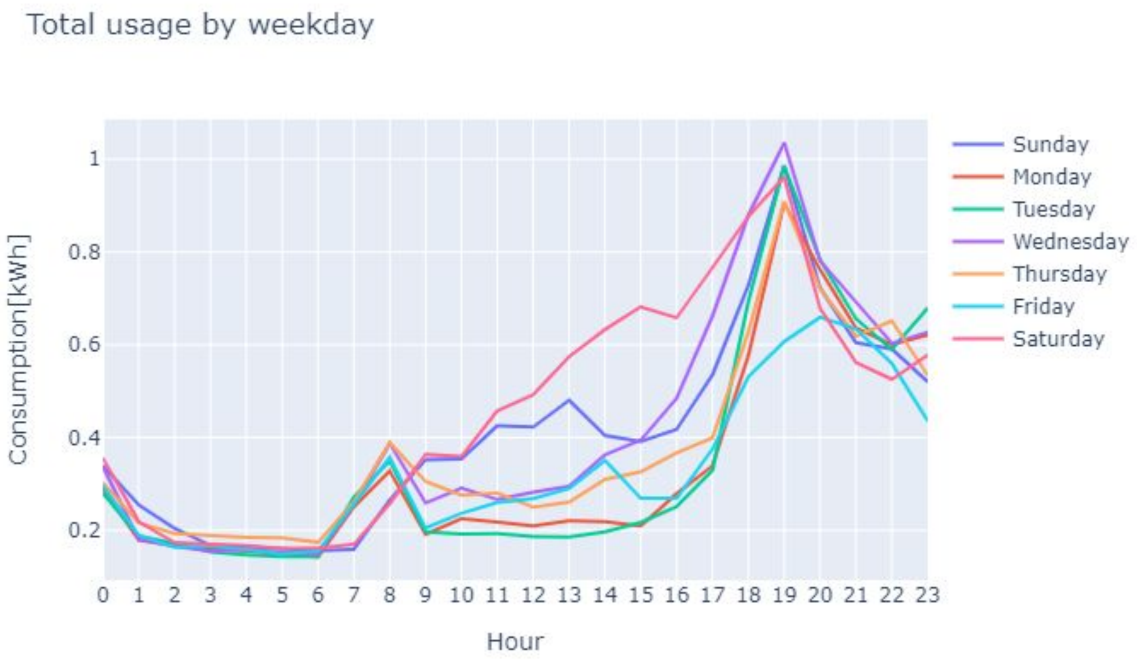


Energy consumption by month, by week

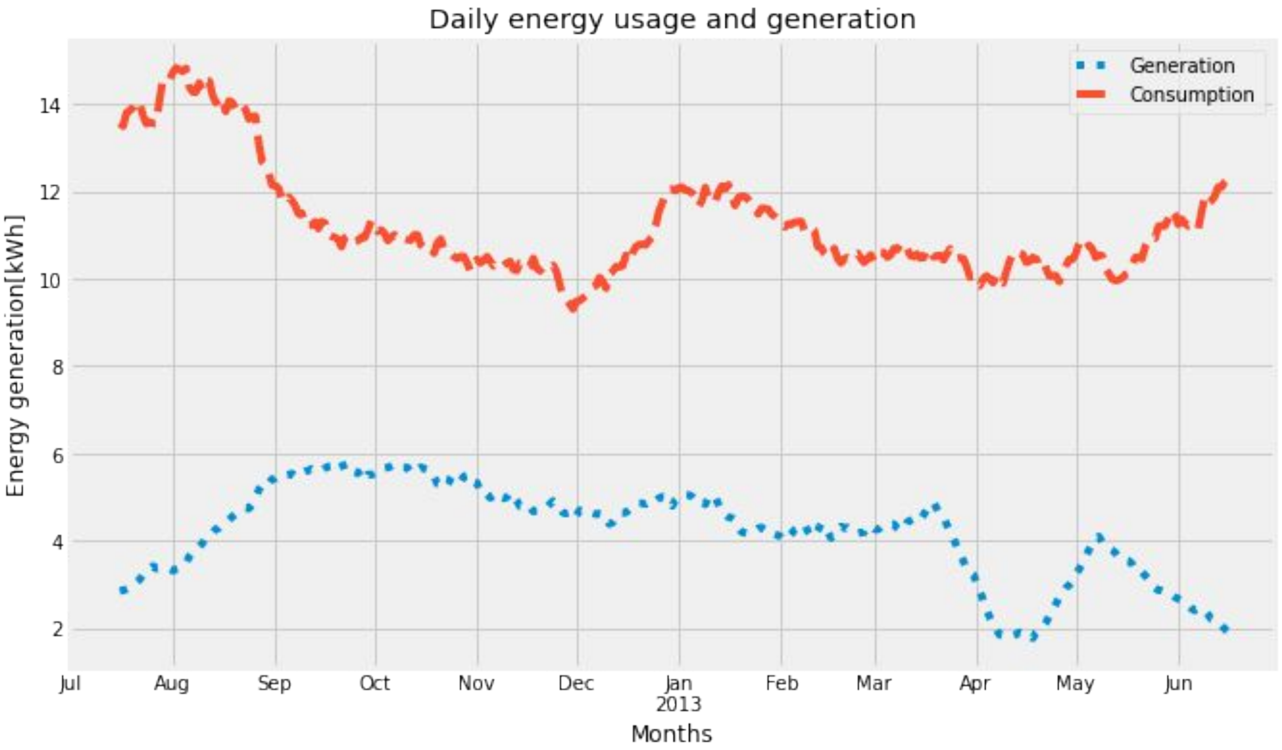




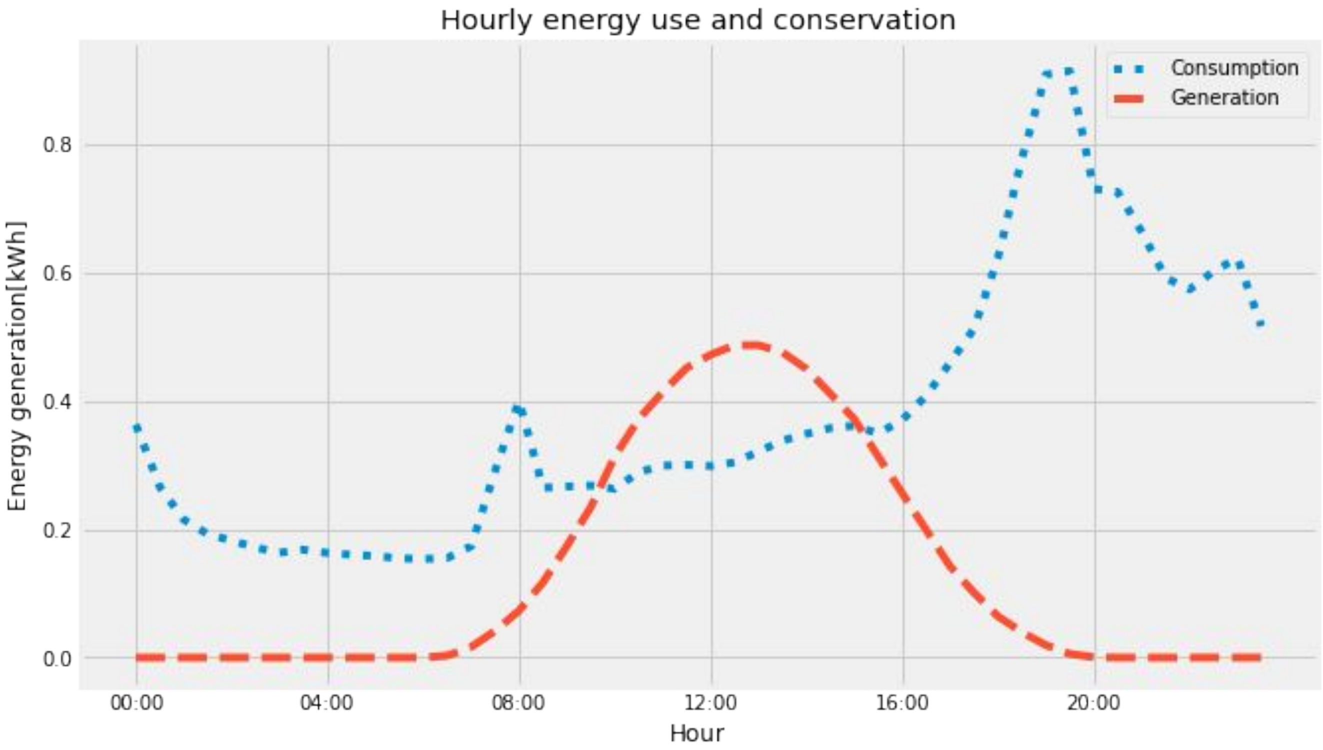
Energy consumption by week and weekend



Daily energy generation and consumption

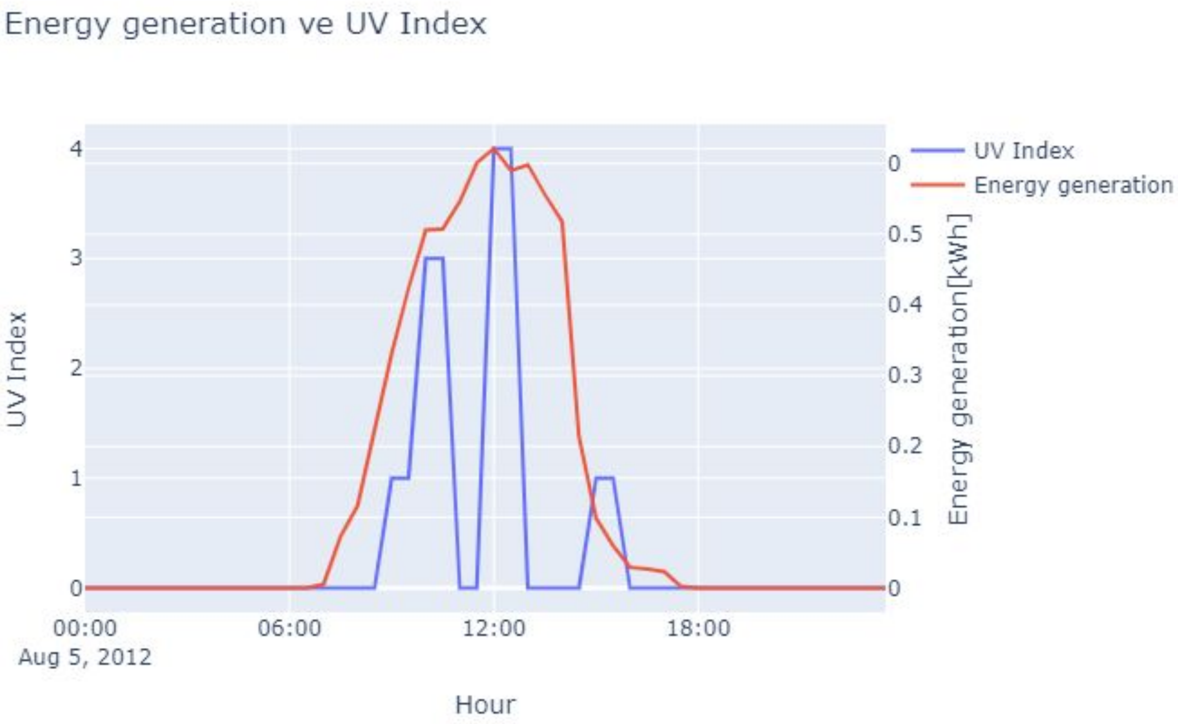


Energy generation and consumption hourly

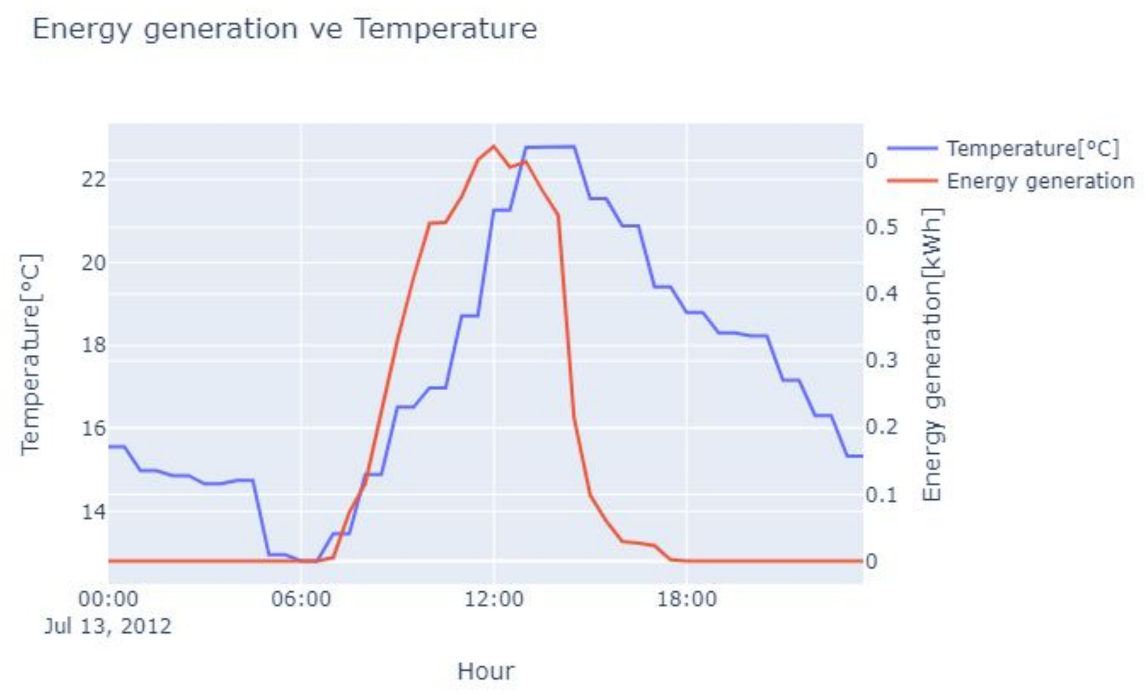


Energy generation and features

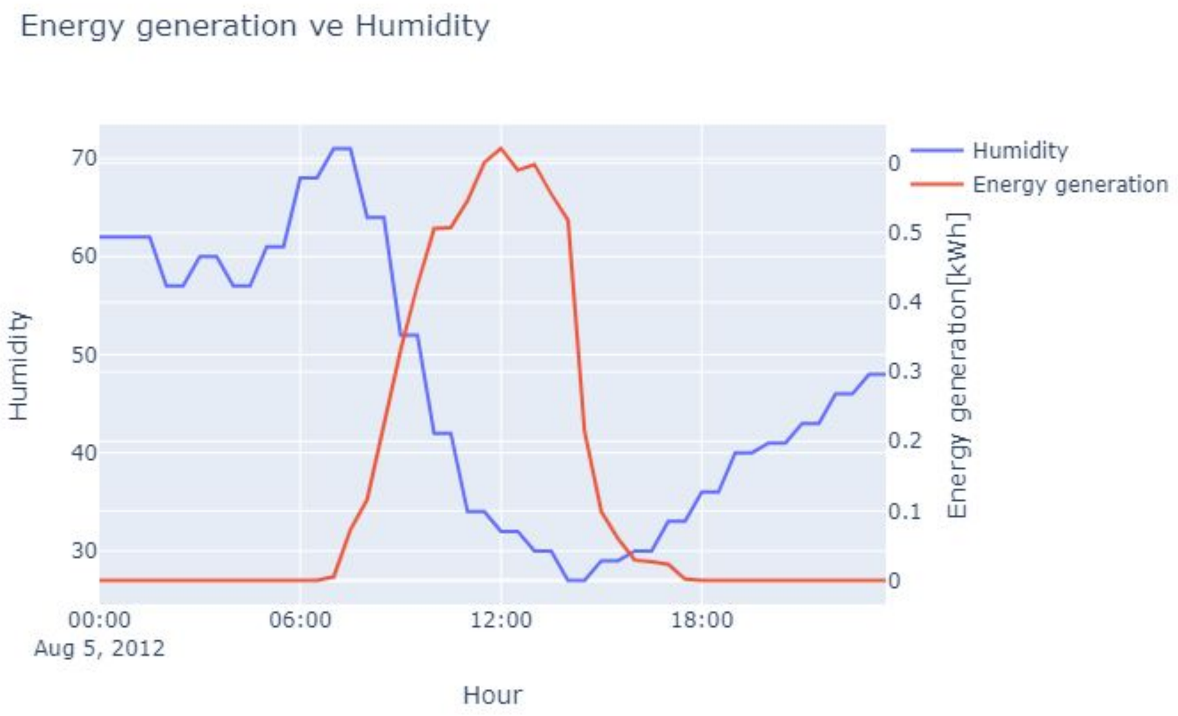
Energy generation vs UV Index



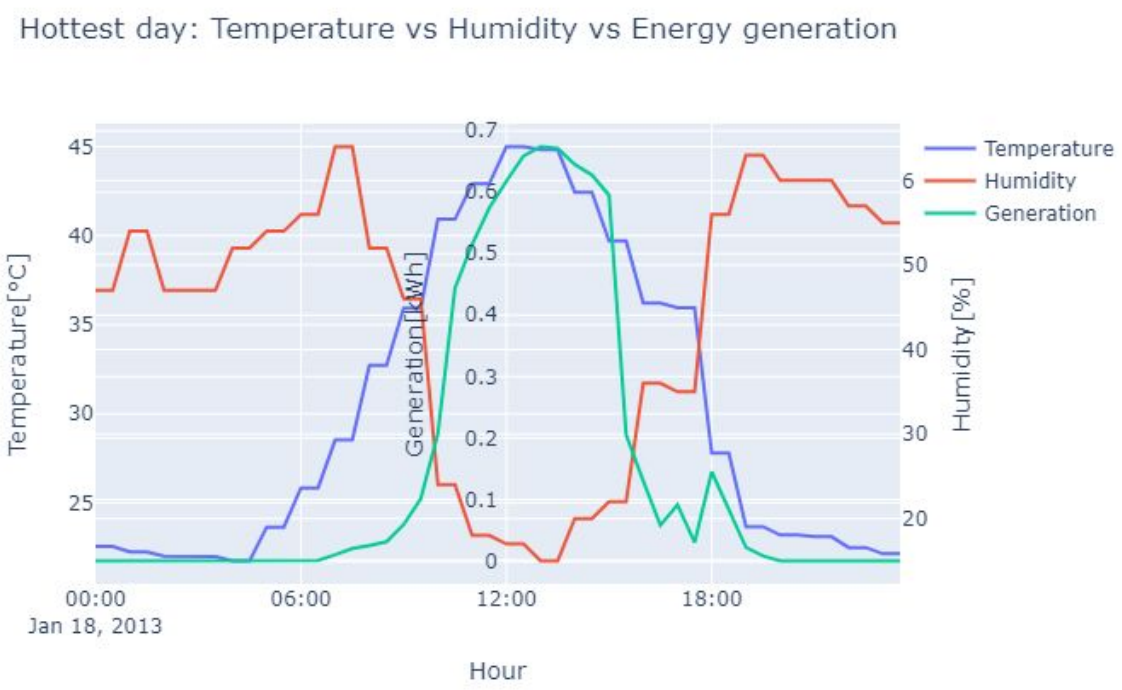
Energy generation vs Temperature



Energy generation vs Humidity



Temperature, humidity and generation

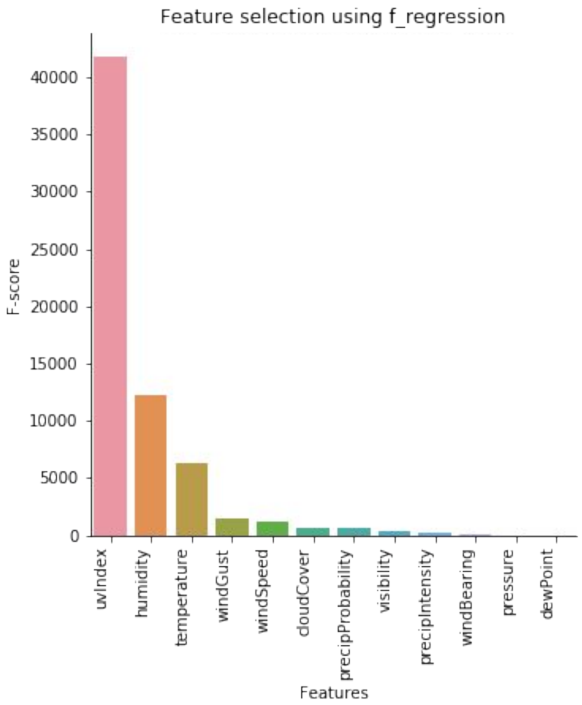


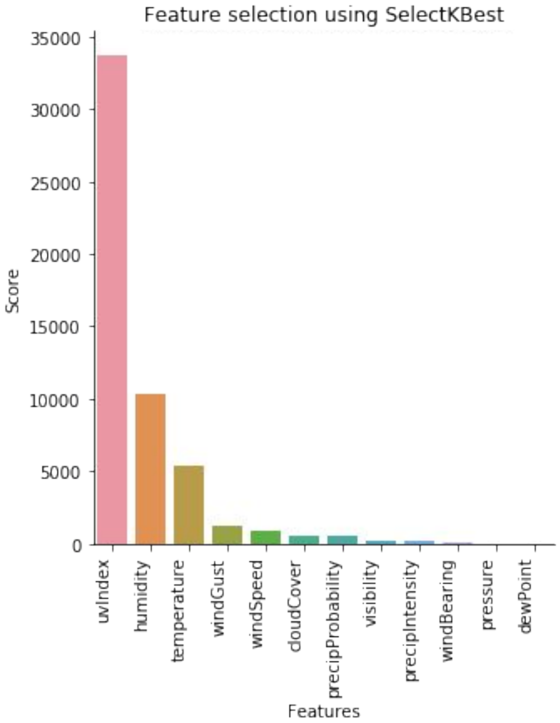
- Modeling on Energy generation

Feature selection

- Most important feature (LR)

- SelectKBest method





Result

Model	Before Feature selection		After Feature selection	
Score	R2 score	MSE	R2 score	MSE
Linear regression	0.58	0.027	0.57	0.027
Random Forest	0.75	0.016	0.69	0.019
XGBoost	0.82	0.012	0.72	0.018

Modeling on Energy consumption

source

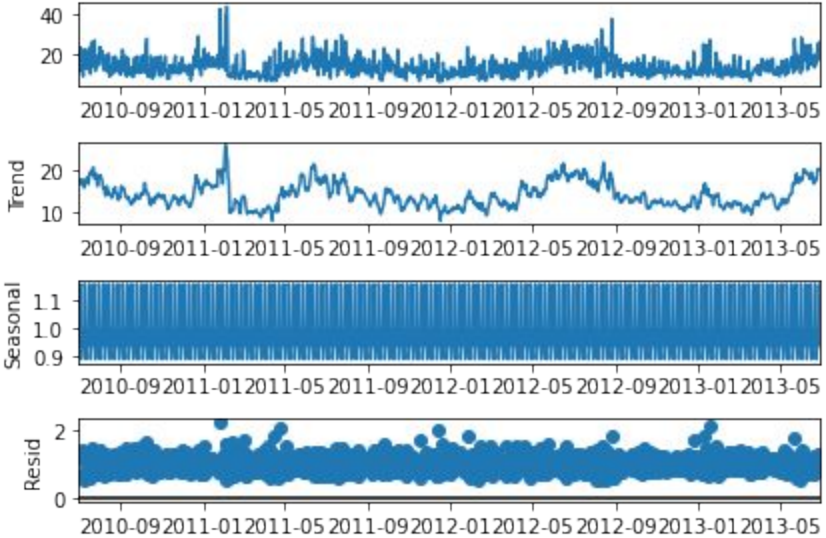
Forecasting Time series

Why time series is special?

- Level,Trend,Seasonality,Noise, Stationarity

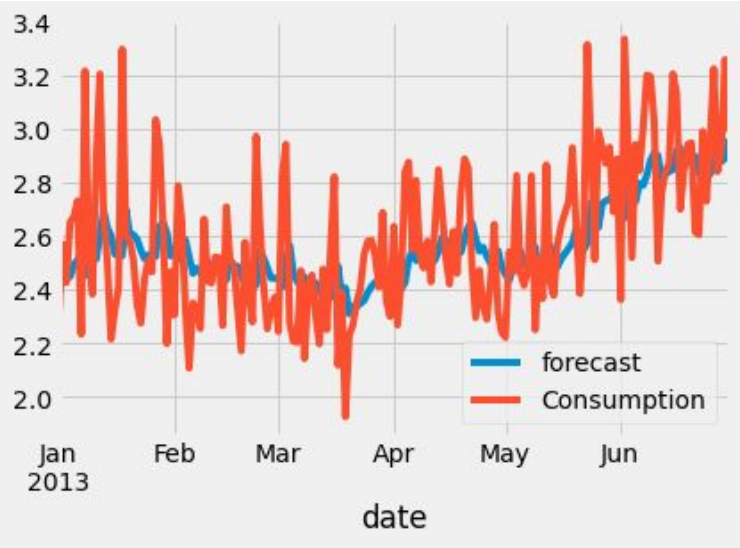
4 models used

- ARIMA
- SARIMAX
- LSTM (deep learning)
- Facebook's Prophet



ARIMA model

ARIMA modeling

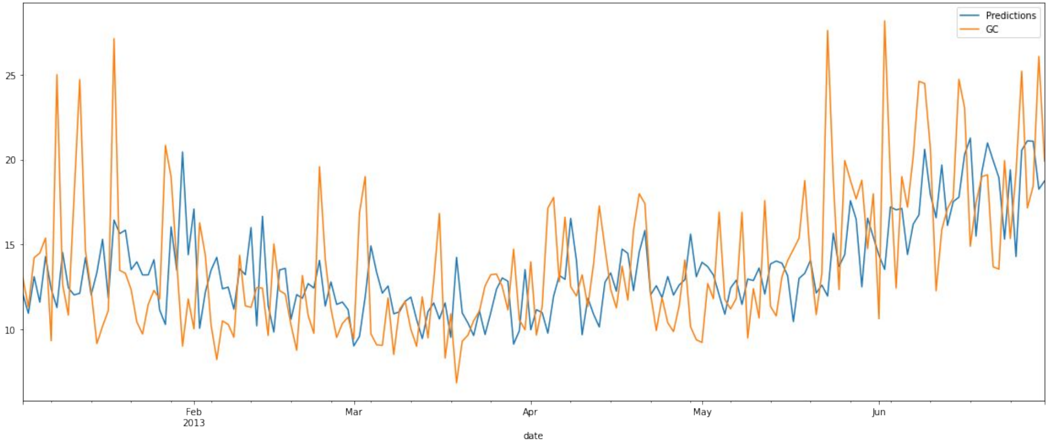


Prediction vs actual consumption

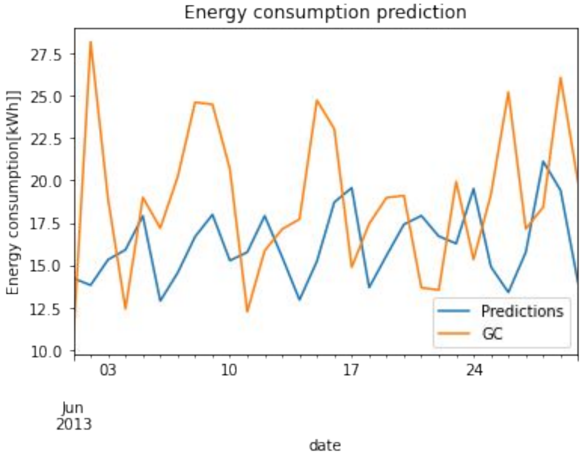
- Find best parameters: auto_arima (functions as GridSeachCV)
    - AR:**p** Auto-regressive aspect of the model, which incorporates past values.
    - Integrate:**d** Integrated part of the model. Effect the amount of differencing to apply to a time series
    - MA:**q** Moving average part of the model
- **Model gets the trend and level. But not much accurate in terms of highest and lowest points.**

SARIMAX model

ARIMA model with SEASONALITY



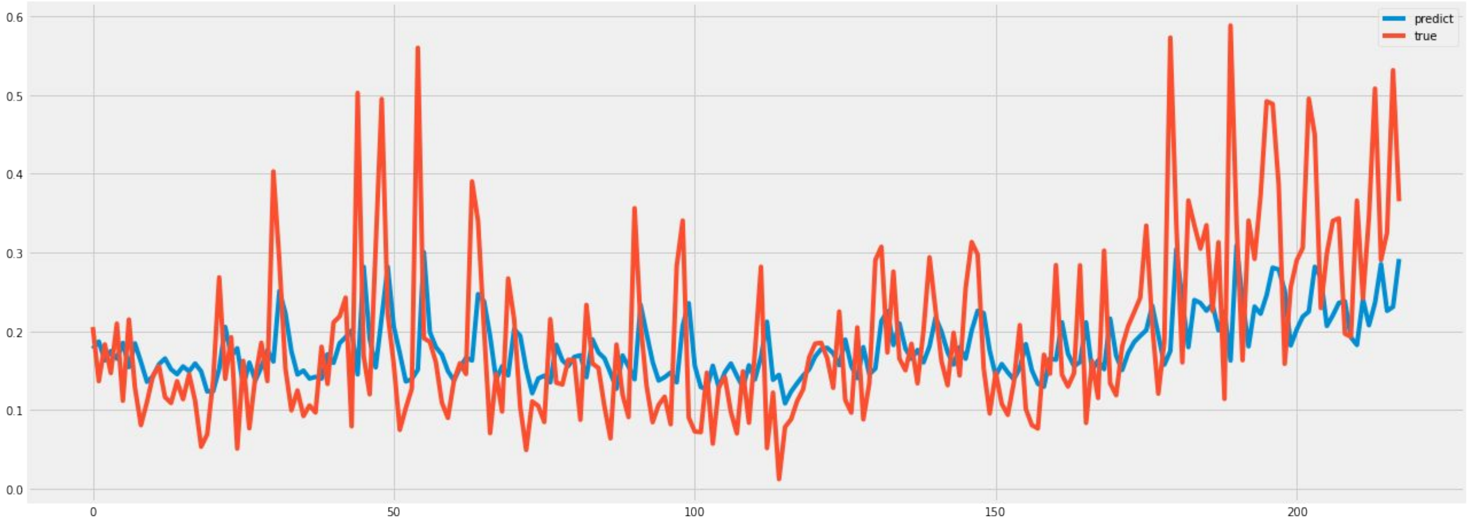
Train set



Test set (prediction)

Naive LSTM model

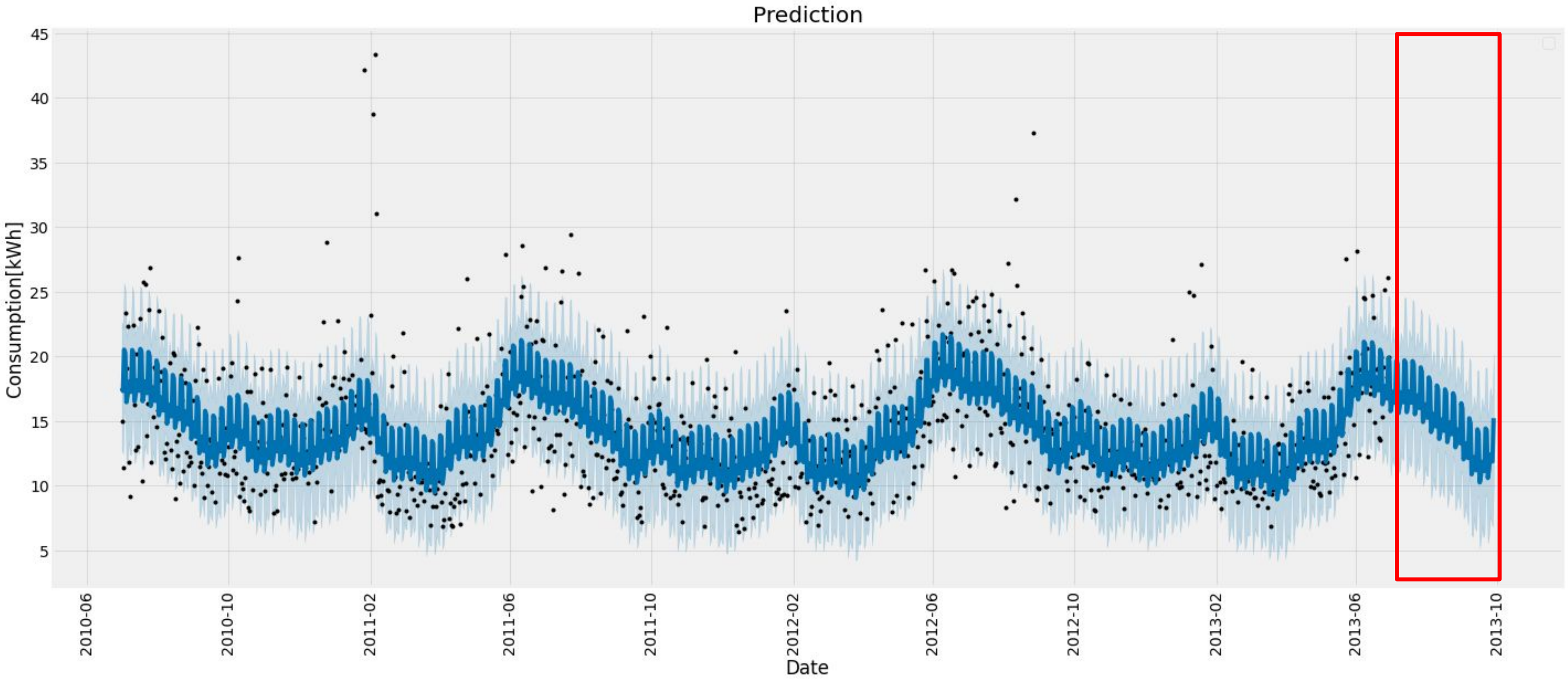
Most simple LSTM model with 2 layers(input 100), 1 Dense layer (output 1), loss function:mae, optimizer:adam.



Result in test set

Facebook Prophet

$$\begin{aligned} y(t) = {} & \text{piecewise\_trend}(t) + \\ & \text{seasonality}(t) + \\ & \boxed{\text{holiday\_effects}(t) +} \\ & \text{i.i.d. noise} \end{aligned}$$



Prediction

Modèles et Résultats

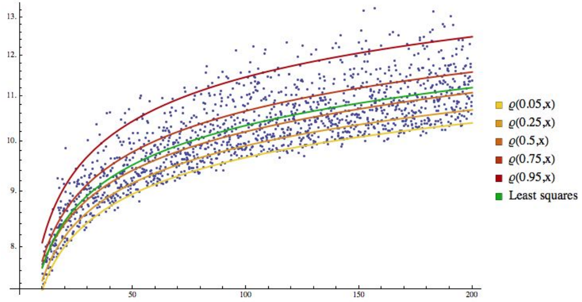
Modèle	RMSE[kWh]
ARIMA	5.4
SARIMAX	4.1
LSTM	3.8
Prophet	5.5

Conclusion & Perspectives

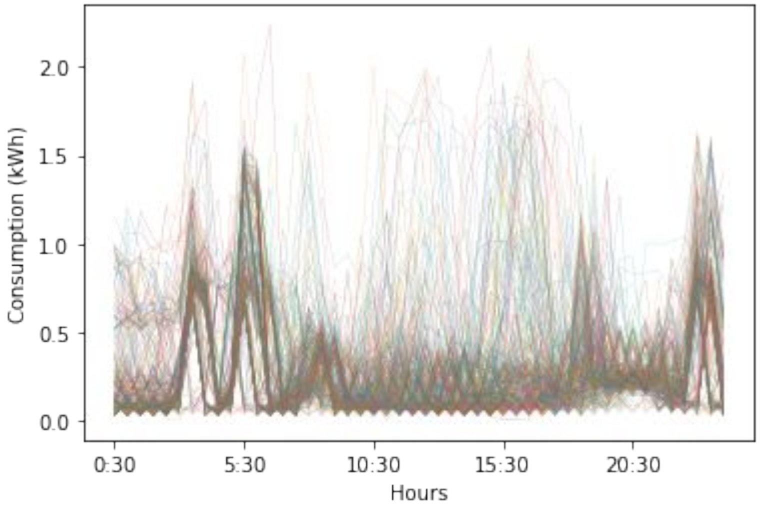
Conclusion

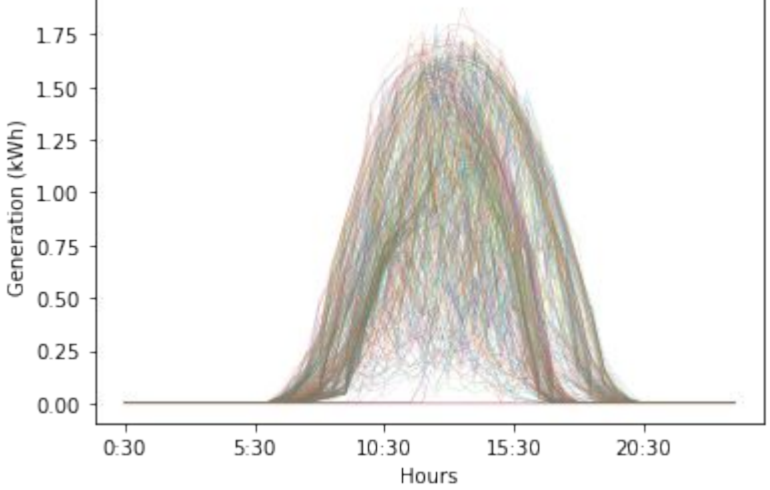
- Obtain more data to make our predictions better
- Radiation (UV Index), temperature and humidity seem to be important parameters for better generating solar energy
- For power generation, the XGBoost model performs best (score 0.82 R2)
- For power consumption, the LSTM model performs best (3.8 [kWh] RMSE).

Future work



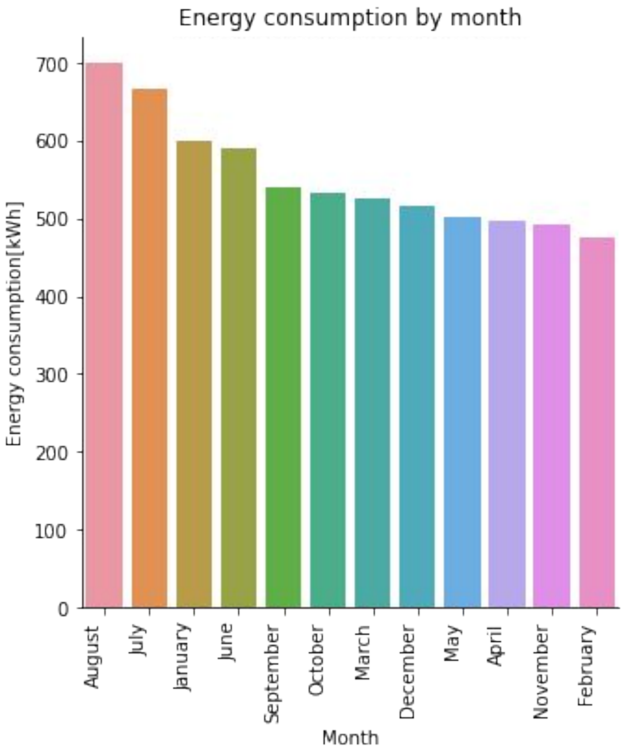
- Use quantile regression K-near neighbors metho energy consumption and generation
- Feature selection for energy consumption prediction

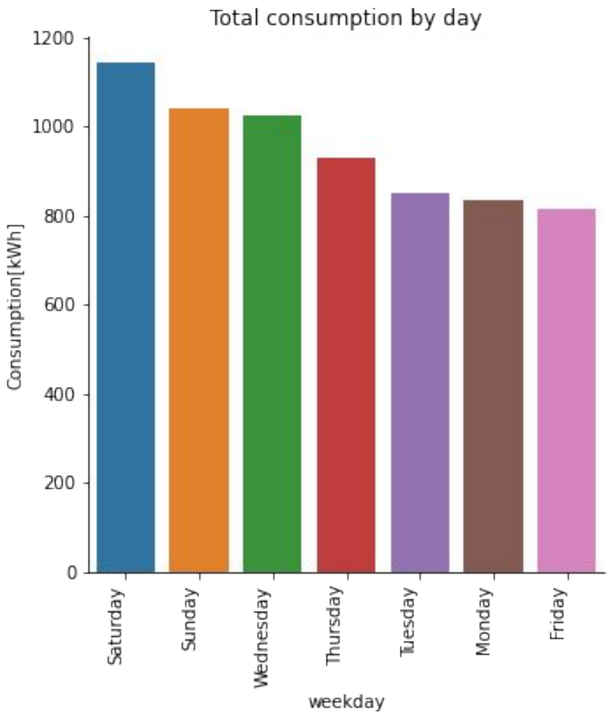




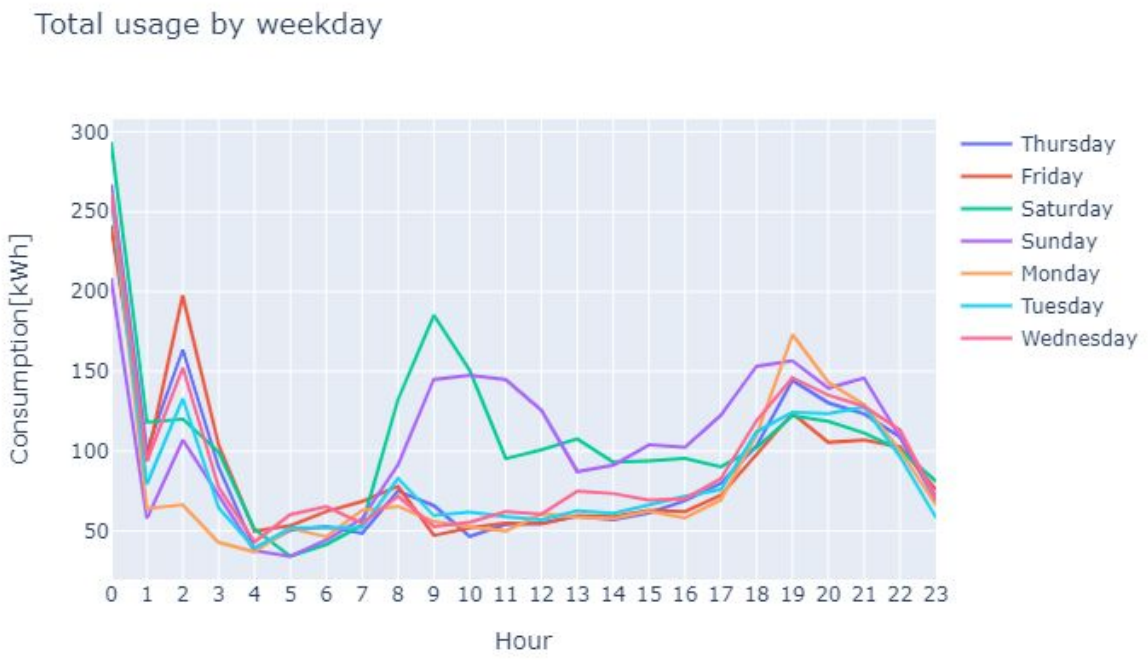
Additional materials

Energy consumption by month, by week

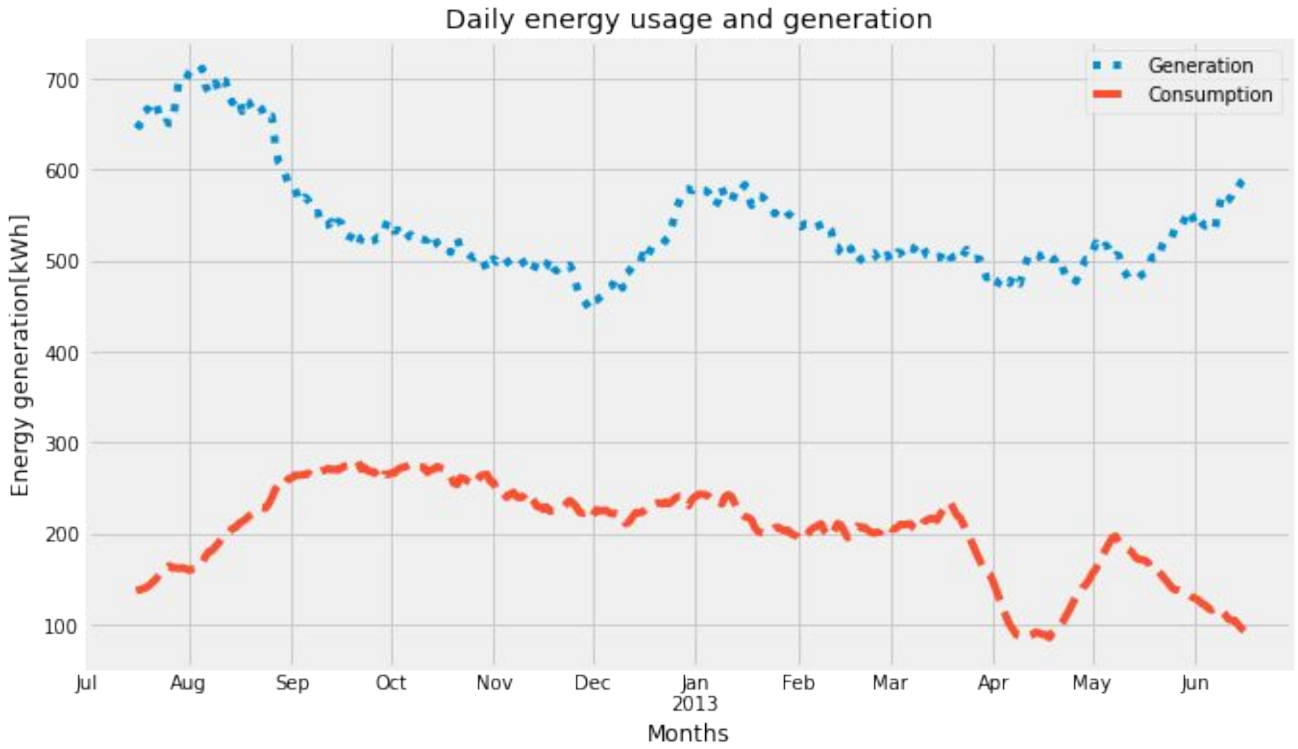




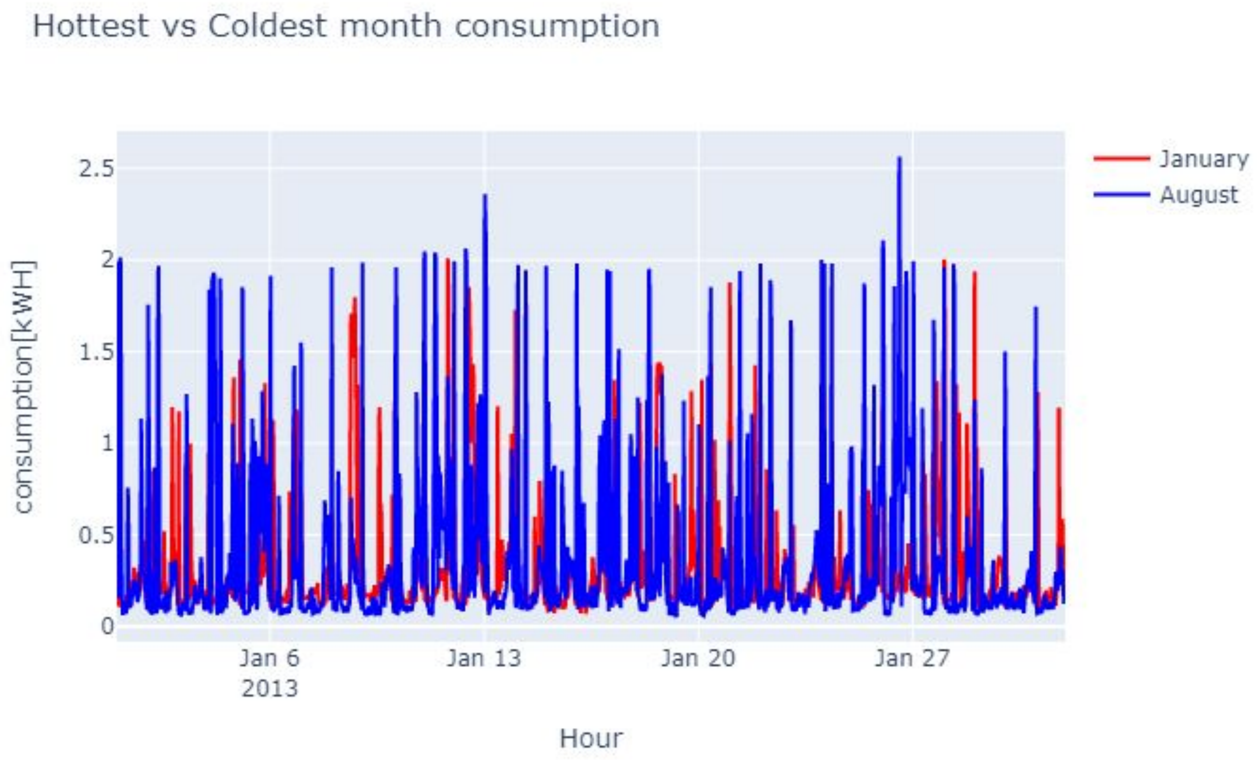
Energy consumption by week and weekend



Daily energy generation and consumption



Hottest vs Coldest month consumption



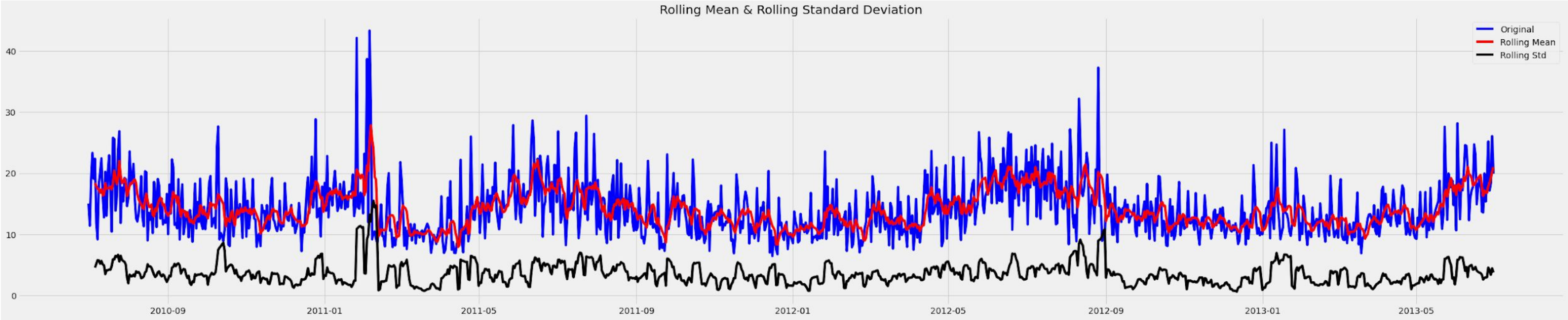
ARIMA model

Important parameters:

AR:**p** Auto-regressive aspect of the model, which incorporates past values.

Integrate:**d** Integrated part of the model. Effect the amount of differencing to apply to a time series

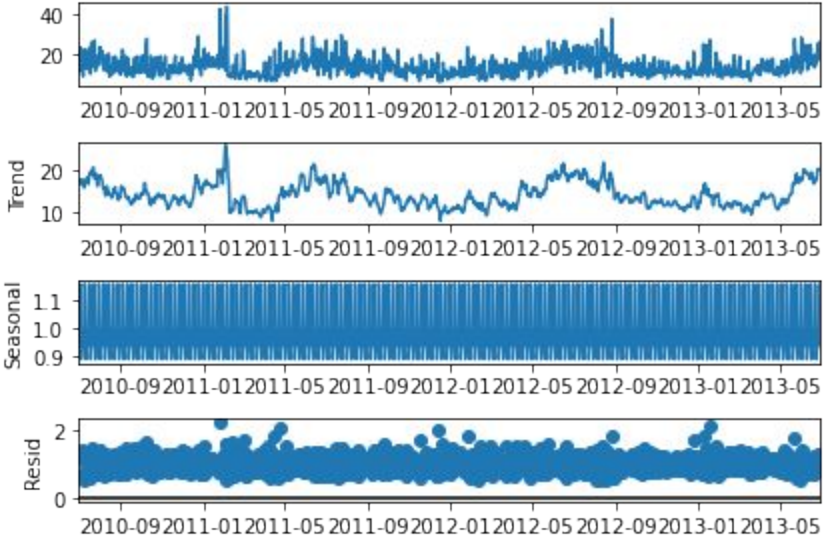
MA:**q** Moving average part of the model



Forecasting Time series

Why time series is special?

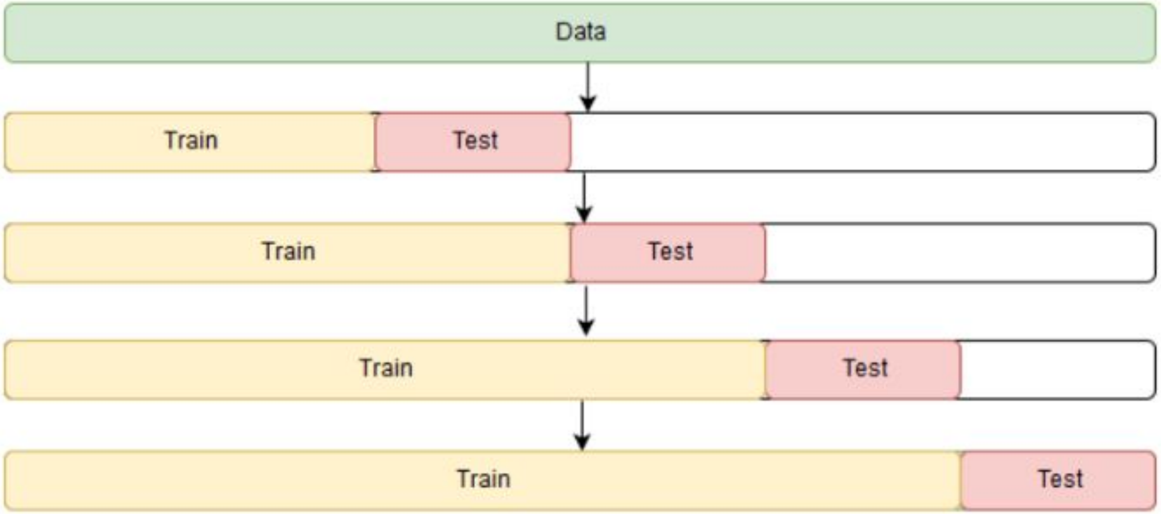
- **Level**: The baseline
- **Trend**: Behavior of the series over time.
- **Seasonality**: Repeating patterns or cycles of behavior over time.
- **Noise**: Variability in the observations
- **Stationarity**: A time series is stationary when the mean, variance, and autocorrelation are constant over time



ARIMA model

ARIMA model

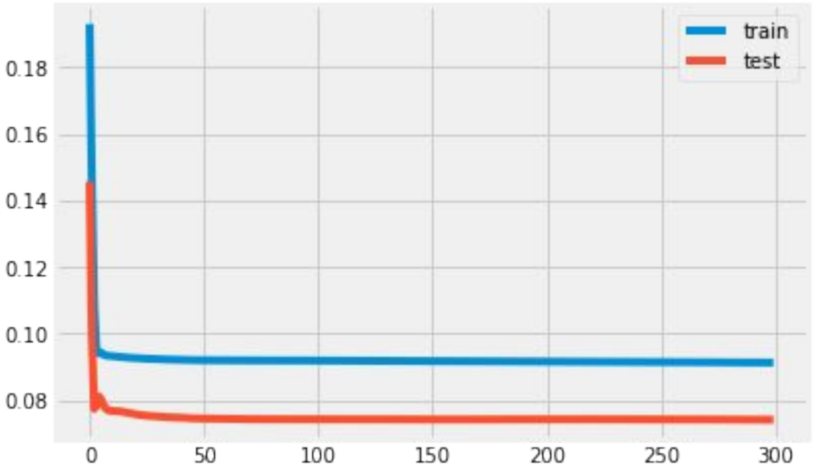
Find best parameters: auto_arima (GridSeachCV in arima)



Window rolling technic

Naive LSTM model

Most simple LSTM model with 2 layers(input 100), 1 Dense layer (output 1), loss function:mae, optimizer:adam.



Loss in train, test set

Facebook Prophet

We can see more detail in Prophet result!

	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper	additive_terms	additive_terms_lower	additive_terms_upper	daily	daily_lower	daily_upper
0	2013-01-01	14.449945	6.785332	16.501677	14.449945	14.449945	-2.667754	-2.667754	-2.667754	-0.358377	-0.358377	-0.358377

weekly	weekly_lower	weekly_upper	yearly	yearly_lower	yearly_upper	multiplicative_terms	multiplicative_terms_lower	multiplicative_terms_upper	yhat
-1.118643	-1.118643	-1.118643	-1.190734	-1.190734	-1.190734	0.0	0.0	0.0	11.782191

	horizon	mse	rmse	mae	mape	mdape	coverage
0	36 days 00:00:00	22.076445	4.698558	3.886572	0.298633	0.228618	0.700935
1	36 days 12:00:00	22.211953	4.712956	3.906536	0.300074	0.233749	0.700935
2	37 days 00:00:00	22.504910	4.743934	3.937652	0.300372	0.234613	0.691589
3	37 days 12:00:00	22.054008	4.696170	3.894398	0.299203	0.233749	0.696262
4	38 days 00:00:00	21.926405	4.682564	3.889139	0.298848	0.234613	0.691589

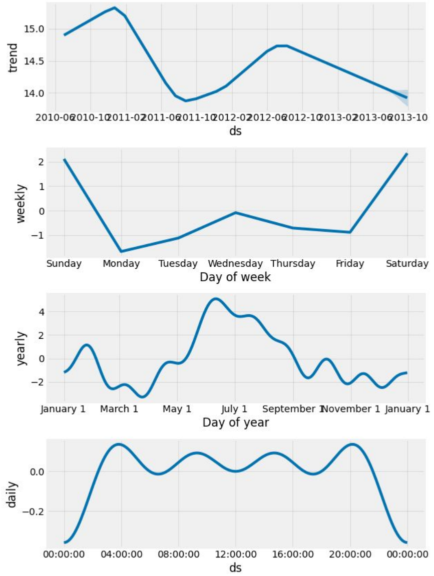
Prophet result

Facebook Prophet

Prophet uses models trend, seasonality and HOLIDAY effects, and irreducible error.

$$\begin{aligned} y(t) = {} & \text{piecewise\_trend}(t) + \\ & \text{seasonality}(t) + \\ & \boxed{\text{holiday\_effects}(t) +} \\ & \text{i.i.d. noise} \end{aligned}$$

Prophet



Decomposition