

Challenge Data

Historical consumption regression for electricity supply pricing

by BCM Energy
currently running

[Introduction](#)[Public ranking](#)[Academic ranking](#)[Intermediate academic ranking](#)[Final ranking](#)[Submissions](#)[Submit a solution](#)

Description



Dates

From Jan. 1, 2019 to Jan. 1, 2020

Challenge context

Planète OUI has to offer green electricity supply with prices adapted to the consumption profiles of its clients. The electricity is indeed bought on the electricity markets, where prices are highly variable during a day. A site with a high share of consumption when prices are low, for example at mid-afternoon, will be supplied at lower costs than a site consuming electricity mainly during peak price intervals. The consumption profile of an installation has then to be appraised to compute the best estimation of supply tariffs, allowing Planète OUI to meet its running costs while supplying its client at the best price. Sites with a metering power larger than 250 kVA are subject to a precise analysis methodology. Most sites are characterized by a consumption varying strongly with temperature because of electrical heating systems. French national electricity consumption is the most temperature-sensitive in Europe, increasing by 2.4 GW per lost degree – the output of 2 or 3 nuclear reactors. Nonetheless, industrial installations are specific cases, because their consumption might be highly related to non-thermosensitive uses, for instance chemical or metallurgical processes. For each site, the objective is namely to analyze thermosensitivity uses. This, however, is not enough to determine precisely the consumption profiles, other factors affecting consumption, for example annual, weekly and daily seasonalities. When Planète OUI prepares electricity supply offer, it receives from the potential client historical consumption data. These profiles are combined with

Public metric

Custom metric given by the provider.

[Download metric file](#)

Course

42 challenge

by Pr. Olivier Crouzet from 42

No report required.

Files

[x_train](#)

input data of the training set

[y_train](#)

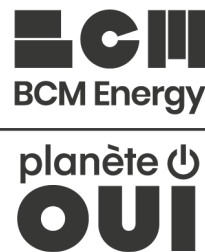
output data of the training set

[x_test](#)

input data of the testing set

The challenge provider

electricity prices simulations to compute a distribution of supply costs in €/MWh. A given percentile is then used to cover supply costs for a wide range of price scenarios. However, the client's data is often incomplete and spread over a relatively short period which is rarely longer than a year. In order to get a more precise supply pricing, the goal of Planète OUI is to develop a machine learning model where one or several years of extrapolated consumption data are rebuilt from a single year of measured data supplied by the client. These extrapolated profiles will be combined with electricity prices as well, in order to get a larger data set of analysis.



BCM Energy is a start-up based in Lyon and created in 2015. BCM operates on the whole value chain of renewable electricity, from production assets management on electricity markets (Epexspot, EEX) to electricity supply to final consumers through the brand Planète OUI: - BCM Energy is active continuously on the electricity markets, - BCM Energy is responsible of its own balancing perimeter, - BCM Energy also manages capacity certificates within its own certification perimeter and guarantees of origin in its register. The diverse priority areas of growth are supported by a trading team with more than fifteen years of experience of the various electricity markets, developing state-of-the-art financial analysis modeling. Planète OUI, created in 2007, is one of the first French green electricity supplier. The company supplies tens of thousands of homes and professionals and covers more than 95 % of metropolitan France (Enedis network). Planète OUI promotes an ecology constructive and made available for all. It has integrated BCM Energy's perimeter in 2017.

[PROVIDER WEBSITE](#)

Challenge goals

The goal of the challenge is to predict, based on the analysis of the correlation of a year of consumption and weather training data, the electricity consumption of two given sites for a test year. In operational conditions, the new consumption profiles would be integrated to electricity supply pricing analysis.

Data description

Data description

The entry data of the model to be developed is composed of the following columns; data being recorded on an hourly basis:

"ID": Data point ID;
 "timestamp": Complete timestamps with year, month, day and hour, in local time (CET and CEST);
 "temp1", "temp2", "mean_nationaltemp": Local and mean national temperatures (°C);
 "humidity1", "humidity2": Local relative humidities (%);
 "loc1", "loc2", "locsecondary1", "locsecondary2", "locsecondary3": the coordinates of the studied and secondary sites, in decimal degrees and of the form (latitude, longitude).
 "consumptionsecondary1", "consumptionsecondary2", "consumptionsecondary3": the consumption data of three secondary sites, whose correlations with studied sites may be of use (kWh). Indeed, the two studied sites and the three secondary sites are used for the same purposes; The output data of the model to be developed takes the following form:
 "ID": Data point ID;
 "consumption1", "consumption2": the consumption data of the two studied sites (kWh).
 Relative humidities are provided with temperature data because they represent variables of importance for electricity consumption: humidity indeed strongly influences thermal comfort. To replicate operational conditions, some temperature and humidity data points will be missing. The imputation method must be carefully considered. The "consumptionsecondary*i*" variables are the consumption data of several sites with metering power higher than 250 kVA of the Planète OUIs portfolio. This correlation of the various sites consumptions shall be studied to precise data completion or interpolation. Timestamps may be expressed as month or day of year, day of week and hours, to study the impact of annual, weekly and daily seasonalities. Particular attention should be paid to national holidays processing.

Benchmark description

Persistency is a fast and relatively accurate benchmark. It consists in estimate that the hourly consumption of the y+1 year is equal to that of the y one, shifting data relative to weekends and public holidays. For instance, the benchmark method would consider the hourly consumption of Saturday, August 4, 2018 is equal to that of Saturday, August 5, 2017.

[I want to cancel my participation](#)



[about](#) | [team](#) | [terms of use](#) | [legal notice and privacy policy](#)