FACULTÉ DES SCIENCES APPLIQUÉES
Ecole d'ingénieurs et d'informaticiens

PROJ0016-1
BIG DATA PROJECT

Will EU achieves its goals of CO_2 reduction by 2030?

MILESTONE 2

VIESLET Thomas - s153205
GIRINEZA Guy - s144377
DELVOYE Benjamin - s154317

Master 1 Data science Engineering
Faculty of Applied Sciences

Academic years 2019-2020

1 Time serie forecasting on CO_2 emissions for each sector

As part of the action adopted by the European Council in October 2014, the renewable energy and energy efficiency targets have been revised upwards in 2018. The European Council is aiming for an average 40% reduction in CO_2 emissions in 2030 compared to 1990 levels. Belgium, for its part, must reduce its CO_2 emissions by 35%[1]. By forcasting we want to check whether Belgium will achieve the objectives set by the Council of Europe. Indeed, we make the hypothesis that Belgium will continue on its latest trend. To do this, we are going to check whether the sum of the CO2 emissions of each sector corresponds to a decrease of 35% compared to 1990 according to each model. Total CO_2 emissions in 1990 amounted to 162.46 Mt.

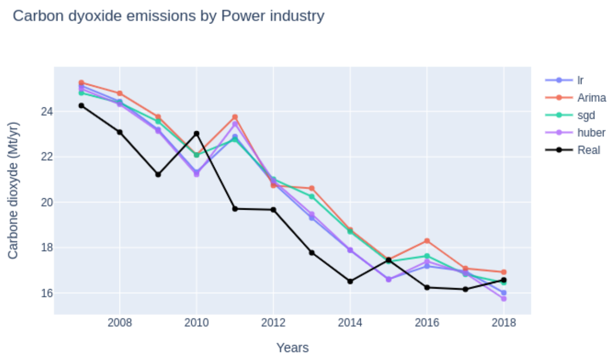
As a starting model, we have chosen to use linear models such as a simple linear regressor, an Huber regressor and a linear regression with stochastic gradient descent. We also applied an integrated autoregressive moving average (ARIMA). We have fetched the data of belgian CO_2 emission for each main emission sector from 1970 and 2018 [2].

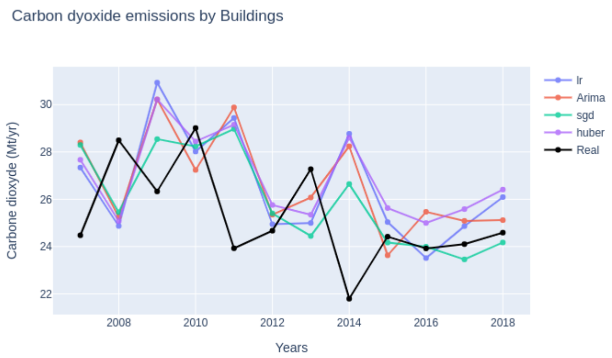
1.1 Evaluation of models

We have compute the mean absolute error for each model in each sector and ranked it in the following tabular.

Industry	Buildings	Transport	Other industrial combustion	Other sectors
Huber: 1.43	Lr: 2.43	Arima: 0.53	Arima: 0.98	Arima: 0.46
Lr: 1.51	Sgd: 2.44	Sgd: 0.54	Huber: 1.12	Lr: 0.53
Arima: 1.65	Huber: 2.45	Lr: 0.61	Lr: 1.13	Huber: 0.59
Sgd: 2.52	Arima: 2.58	Huber: 0.64	Sgd: 1.85	Sgd: 0.65

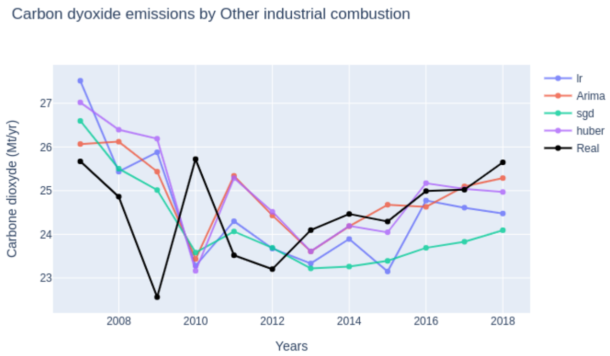
This tabular shows no clear winner model in all the sectors.

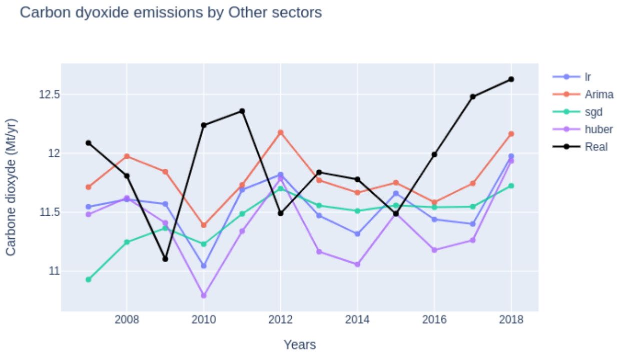




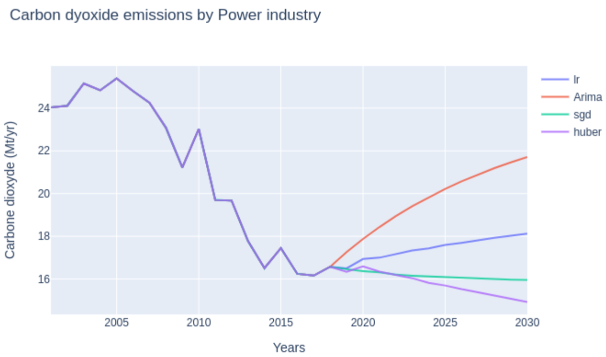
[1] https://www.europarl.europa.eu/news/en/headlines/society/20180208STO97442/cutting-eu-greenhouse-gas-emissions-national-targets-for-2030

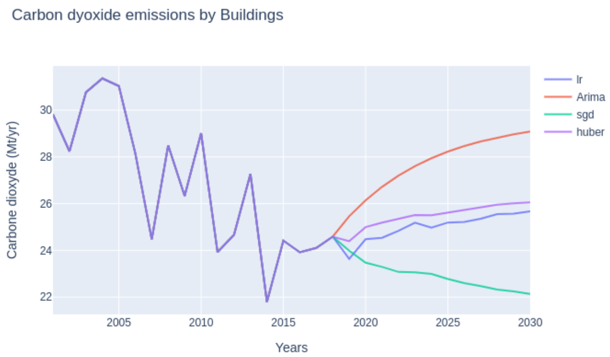
[2] https://edgar.jrc.ec.europa.eu/overview.php?v=booklet2019

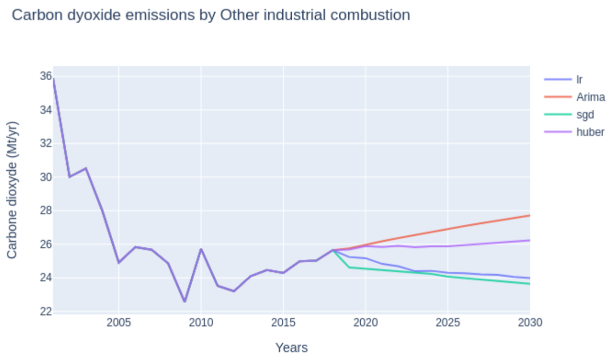


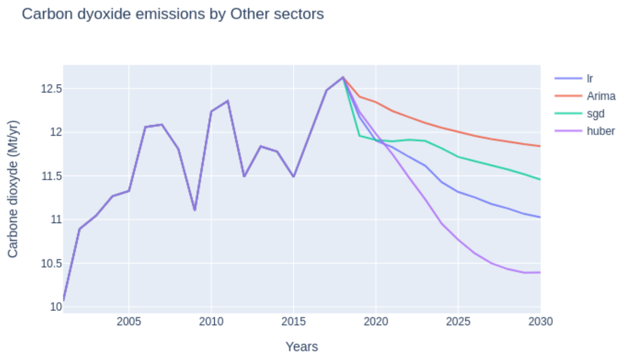


1.2 Forecasting









Model	Total sum (Mt) in 2030	% (compared to 1990)
lr	105.94	34.8
ARIMA	116.1	28.6
SGD	101.72	37.4
Huber	104.64	35.6

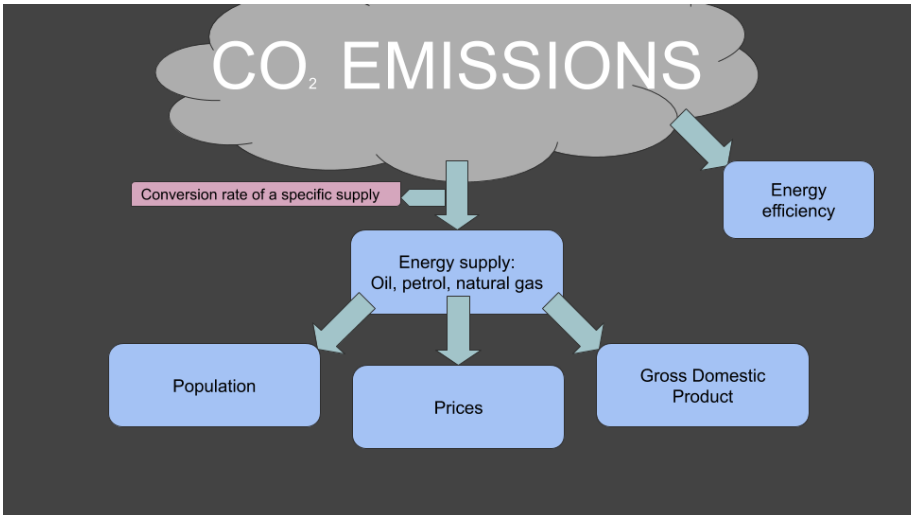
Those results are based on the trend of the past few years. It is assumed that the variables are independent, so emissions from each sector do not affect other sectors. This is obviously a too simplistic assumption. We can see that if Belgium goes on with this trend, it will match the objectives set by the Concil

2 What's ahead?

For the next milestone our goal would be to get a similar interface as find on this website[3]. To achieve this we will first pick variables to include in our analysis. Then we will analyse the set of data to highlight structure of correlations.

Once this is done, we will forecast most independent variables and finally estimate CO_2 consumption based on predictions and correlations found earlier.

We present below a first shot of a chart explaining links between variables.



[3]https://en-roads.climateinteractive.org/scenario.html

3 First version of the model

3.1 Variables selection and correlation analysis

In order to predict future value for CO_2 emissions it is important to find which parameters impact CO_2 the most and how.

For this milestone we focus on the analysis of primary energy consumption in Belgium and their relations trough time. We gathered data from 1965 to 2015 on consumption about oil, petrol and gas. We choose to lean on these 3 quantities because we made the assumption that they represent the main portion of CO_2 emissions (as in these consumption also consider the production of electricity from these energies).

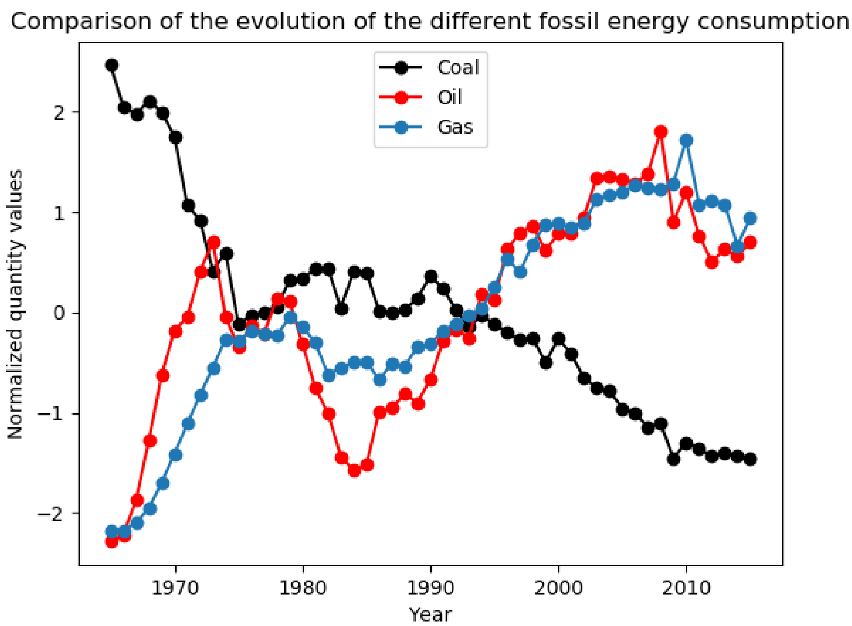
However, other factors will be taken into account later on. For example, the overall electricity consumption will impact these consumption (and so, CO_2) as some oil/gas/coal consuming products can be substitute by other electricity consuming products (oil car → electrical car, coal boiler → heat-pump, ..). A bunch of factors impact these quantities and we will have to consider them on due time but we decided that it was a relevant first consideration.

We have computed the correlations between the different energies consumption and obtained the followings:

	Coal	Oil	Gas
Coal	1	-0.74	-0.93
Oil	-0.74	1	0.84
Gas	-0.93	0.84	1

These obtained values can be interpreted. We observe a strong negative correlation between oil and coal, a very strong negative correlation between gas and coal and a strong positive correlation between gas and oil.

Intuitively, gas and oil can be seen as substitute to coal while they seem to follow the same growth. In order to verify these assumptions, we computed the normalised time series of each energy consumption on the same graph to compare. We obtained the following:



The above plot clearly shows a joint evolution of oil and gas in opposition of the coal consumption. However, these results need to be considered carefully. Indeed, the decrease of coal consumption is not necessarily a cause of the growth of oil and gas consumption nor the other way around. These two phenomena can simply be caused by mostly the same time dependant variables as our 3 variables are themselves time dependant.

Although correlation doesn't mean causality specially in our case, we can logically understand that a decrease of coal consumption indeed provide a bigger share of the market to substitute energies like oil and gas and therefore this should be taken into consideration in our latter development. Nevertheless, no logical cause-consequence relation can be found between the growth of oil and gas consumption (at least by us) which makes their correlation dangerous to handle.

3.2 Next Milestone

We will perform analysis on impacts of more variables such as population, prices and GDP on energy supplies consumption. Then we will define different scenarios about those variables with simple models and use the analysis to estimate the energy supply consumption according to defined scenarios.
To finish we would need information about CO_2 release for each type of supply.

4 Bibliography

- `https://statbel.fgov.be/`
- `http://perspective.usherbrooke.ca/bilan/tend/BEL/fr`