Project Report for 4GMM Internship

Simulating new wind and solar power profiles for a power electricity mix supporting the demand in France in 2050

Author

Amandine Capitaine

Tutors

Aude Rondepierre

Charles Dossal

Researchers at the Institut of Mathematics of Toulouse (IMT)

Department of Applied Mathematics,
National Institute of Applied Sciences (INSA) of Toulouse

October 9, 2022

Contents

Acknowledgments

Fisrt, I would like to thank my tutors Aude Rondepierre and Charles Dossal, who, first of all, proposed me a subject in agreement with the environment theme that I wanted to develop. I would also like to thank them for having made sure that the scientific tools I used during my internship were more statistical and optimization oriented. Furthermore, I am grateful for their help throughout my research internship and for their advice that guided my internship.

Then, I wanted to acknowledge Behrang Shirizadeh, Quentin Perrier et Philippe Qurion for writing their scientific article and for publishing it in opensource to allow other scientific to read it for free and maybe continue the project.

I also thank Clément Lejeune as well as Frnak Kouassi for helping me on the question of using time series for modelling the production of renewable energies.

Finally, I would like to thank the fifth year students who worked on this projet because their reports with a lot of explications concerning the article and the code helped me to start my intership.

Introduction

In a context of climate change, the question of the management of electricity production in France has multiple stakes and is inevitable in view of the Paris agreements. Indeed, the choice concerning the composition of the energy mix in the coming decades, i.e. the technologies used for the production of electricity, will imply a reorganisation of the electricity network on the one hand, and a launch in the construction of parks, whether nuclear, wind or solar on the other.

To help the actors of the energy transition, numerous scenarios of electricity consumption in 2050 have been developed by different actors of the ecological transition but also of the management of the French electricity network. Using these scenarios, a team of researchers has created a model to propose a possible composition of the mix in the 2050 horizons essentially from renewable energies. The principle of the model is to optimise the cost of the mix while meeting French electricity demand hour by hour according to three different consumption scenarios.

In the context of my research internship, I first of all got to understand the tools at my disposal: the Energy Optimization for Low Emission Systems (EOLES) model. The second part of my internship was devoted to the development of production models for the different renewable generation technologies: solar, onshore and offshore wind. The objective of this work is to test the robustness of the model: can it meet the demand every hour for various weather conditions?

The internship report is structured around 6 sections. First, I will introduce my motivations for this internship and the project in which it is part, then I will present the optimisation of the mix in the EOLES model. In a second part I will approach my internship subject which is the modelling of new production profiles by presenting first the characteristics of these time series and then the different methods used to model them. My last two parts are devoted to the analysis of the new simulations of the EOLES model with these new profiles and to the limits of the model.

1 Context and objectives of the intership

1.1 Internship goals

As part of my fourth year internship, I wanted to seize the opportunity to discover mathematics applied to environmental issues. Indeed, the management of resources in the face of global warming is a very topical subject that can be explored in data science. I wanted to improve my skills in optimization while including the statistical side in the project. The subject that my tutors proposed to me met my expectations.

I chose to do an internship in research to discover the way researchers work and the missions they are given. Having some doubts about the profession of researcher, I wanted to experience its different aspects during this three-month internship to help me in my future training choices.

My tutors, Aude Rondepierre and Charles Dossal are two researchers and teachers at INSA Toulouse. Ms Rondepierre is affiliated with the LAAS laboratory (Laboratory of Analysis and Architecture of Systems) a unit of the CNRS (National Centre for Scientific Research). This unit works on the understanding of complex systems in the fields of aeronautics, space, health, energy or communication networks. His research focuses on various optimisation techniques such as non-smooth, non-convex and constrained optimisation in probability.

Mr Dossal is a researcher at the ITM (Toulouse Institute of Mathematics) which is a Joint Research Unit in the sense that it is composed of researchers, engineers, technicians and PhD students. This laboratory explores various mathematical problems, both theoretical and applied. In research, it also works on the theme of optimisation, in particular on convex optimisation, more particularly on optimisation algorithms and the link with ODEs.

My internship is part of a project that aims to explore the questions raised by the study of Behrang Shirizadeh, Quentin Perrier and Philippe Qurion: "How sensitive are optimal fully renewable power systems to technology cost uncertainty? ". Thus, after having taken in hand the model proposed in the article, I looked into the question of modelling new renewable energy production profiles such as wind and solar. This modelling aims at feeding the database of hourly productions to test the robustness of the model.

1.2 EOLES Model

1.2.1 Global presentation of the ENR project (Renewables Energies)

The project is based on the scientific report of a group of researchers [1]. Published in 2019 by three researchers from the International Research Centre for Environment and Development, the study presents the EOLES (Energy Optimization for Low Emission Systems) model. This model consists in optimising the cost of the French energy mix for 2050, under the assumption that it is composed of renewable production and storage technologies, and under the constraint that electricity production meets the demand hour by hour.

Different factors influence the model simulation such as the consumption scenario in 2050, the wind and solar generation profiles, the flow and capacity constraints associated with each technology.

The scientific paper presents 315 scenarios that differ in the assumptions made about the cost of the technologies because the model is very sensitive to these variations. Thus, each scenario corresponds to a realisation of the model with a cost per technology varying in a chosen price range [Table 3 page 15 [1]].

Several particularities of the model are to be taken into account: we assume that France constitutes a single node and that its production is homogeneous at all points; we have the knowledge of the production and consumption hour by hour over the whole year at the beginning of the optimization of the mix; we do not consider the possibility that the demand is flexible and that the system cannot supply electricity at a given time.

1.2.2 Technologies used and model assumptions

Two categories of technologies are present in the mix: generation and storage.

The 6 generating technologies include onshore and offshore wind, solar photovoltaic, run-of-river and lake hydro, and biogas combined with gas turbines. Wind and solar power represent the largest share of production in the mix: for the year 2006 their contribution to the mix is 90%. These energies are dependent on the weather and operate on a seasonal basis, which makes their production intermittent and all the more complex to manage. The hydraulic turbines for rivers also depend on the flow of the river. Hydroelectric production in lakes is based on the flow of water held in the dam. Biogas is derived from the fermentation of organic matter; its combustion in gas turbines allows electricity to be generated on demand in a renewable way.

The storage technologies of the EOLES model are pumped storage, batteries and methanisation. The principle of pumped storage is to channel water into an elevated storage facility when there is a surplus of electricity production so that hydraulic energy can be used during a shortage of electricity. Lithium batteries are used to store electricity on a short-term basis.

1.2.3 Consumption and production scenarios

- <u>RTE:</u>

  RTE, "Réseau de transport d'électricité", is the French transmission system operator responsible for the public high-voltage electricity transmission network in France. The company has proposed electricity consumption scenarios for 2050 [2], the one used in the EOLES model estimates annual consumption at 580 TWh.
  The replacement of fossil fuels will mean an inevitable increase in electricity consumption. 3 areas that will consume more :

  - Low-carbon hydrogen produced by electrolysis for industrial and heavy transport needs
  - Industry: growth in production and significant electrification of processes
  - Transport: end of thermal vehicle sales for 2035 (strong increase in electric vehicles)

Moreover, no RTE scenario foresees a nuclear phase-out in 2050, as it would require a tenfold increase in the rate of installation of wind turbines and panels, and accepting shortages. The worst case would be to resort to gas-fired power stations, which would multiply emissions and spoil the efforts made. Unless we adapt our way of life and review our consumption in a strong way as proposed by the négaWatt association [3].

- négaWatt:

  The négaWatt association proposes the 2050 scenario with the lowest electricity demand [4]. This strong reduction in electricity demand (272 TWh per year) is due to the reduction of energy consumption.

  The three axes of the scenario are based on:
  - Sobriety in individual and collective energy use
  - Efficiency: reducing the amount of energy needed to satisfy these needs
  - Priority to renewable energies to replace fossil and nuclear energies

  To achieve its objective of halving electricity consumption, the following measures are proposed: the reduction of waste and meat consumption by 50%, double the amount of pastureland (by 2030), halve intensive livestock farming, support for maintaining and switching to organic farming and ban on synthetic inputs.
  Concerning renewable energies, négaWatt proposes to make wind power the first renewable energy in 2050 (19,000 land-based wind turbines / 3,000 maritime). Regarding solar, the association want to couple agriculture and solar farms.
  For fossil fuels and nuclear power négaWatt wants to shut down power plants older than 50 years, train for nuclear conversion and set an energy taxation.

- Ademe:

  The Ecological Transition Agency is a public institution under the supervision of the Ministry of Ecological Transition and the Ministry of Higher Education, Research and Innovation. Ademe's objective is an electricity consumption of 422 TWh per year [5].

  The areas concerned by the proposed ecological transition are the thermal renovation of buildings, the adaptation of transport and regional planning, the energy production, storage and use, the preservation and restoration of ecosystems, the circular economy and the reduced dependence on scarce resources.

2 Optimising the mix for 2050

2.1 Data-set

2.1.1 Input data

All the variables and the parameters of the model are well describe in the section "Appendix 4: the EOLES model" of the article [1]. In this part we briefly describe the main data used. Three types of input data are provided to the model:

- Production profiles:

For each technology, that is to say offshore, onshore, and solar power, we have 18 years of hourly production in GWh. All these profiles, from 2000 to 2017, have been generated thanks renewable.ninja website [6]. This source provides hourly capacity factor between 0 and 1 that characterises the efficiency of production at a given time. These profiles have been created by means of satellite observations and with a reanalysis of the data. Thus we have for each simulation three profiles of capacity factors of size 8760. To obtain the production this profile is multiplied by the installed capacity.

- Consumption profiles:

To run the model we also need the electricity demand curve ; there are 3 scenarios available : RTE, Ademe and NegaWatt which were described previously. They have of course the same size of the production profiles and the unit of measurement is GW. The demand in power of the year 2006 is also provided and will be useful for the results analysis.

- Cost data:

Among the cost data can be found operation and maintenance cost, the fixed ones for each generative and storage technology and the variable only for solar and wind power. An other parameter is the annualised investment cost of each technology. The values of these parameters are calculated in accordance with the scenario trajectories for 2050.

If we run the model without optimising the initialisation data i.e. the resources available at the beginning of the year, we also have the provide the installed capacities of production in GW, the volume of energy storage in GWh and the capacity of storage in GW.

2.1.2 Output data

After the simulation of the model we obtain:

- COST: overall investment cost over the year
- G: electricity generation per technology (hourly)
- STORAGE: electricity entering in a storage technology (hourly)
- STORED: stored energy per technology (hourly)

If we optimise the initialisation data, the following data will also be optimised:

- Q: the installed capacity (what the technology could produce at its maximum); for a storage technology Q is the output flux.
- S: the entrance flux for a storage technology.
- VOLUME: the maximum energy of a storage technology that we can store.

2.2 Model

2.2.1 Objective function

The objective function aims to minimise the cost of the energetic mix composed of the installed capacities and storage volume annuities plus the fixed and variable costs of generative and storage technologies.

$$
\begin{aligned}
COST = \Bigg(& \sum_{tec}[(Q_{tec} - q^{ex}_{tec}) * annuity_{tec}] + \sum_{str}(volume_{str} * annuity^{en}_{str}) \\
& + \sum_{tec}(Q_{tec} * fO\&M_{tec}) + \sum_{str}(S_{str} * (capex^{ch}_{str} + fO\&M^{ch}_{str})) \\
& + \sum_{tec}\sum_{h}(G_{tec,h} * vO\&M_{tec}) \Bigg) / 1000
\end{aligned}
$$

The authors of the article have also developed a simplified model which only optimises the variable costs link to the hourly production of every technology. In fact, the objective function of the smallest model is only the last line of the above equation because the variables volume, installed and charging capacities are fixed in this case.

2.2.2 Constraints

The first constraint is the need to **satisfy demand** every hour, translated into an inequality: the power produced must be greater than the power consumed plus the energy stored.

The model then add **physical constraints**: the electricity generated must not exceed the maximum amount of electricity that can be produced with the installed capacity. For physical reasons we also define maximum of capacity installed for every technology as well as the storage volume.

The **natural constraints** are define for biogas because of the limited resources. Concerning hydroelectricity generated by lakes we have to respect the non-energy operating constraints. This is traduced by a limit of production per month in the model.

Storage constraints are implemented to avoid an ending year without any stored energy: we force the model to have the same stock as at the beginning of the year. Moreover, the model has constraints related to the storage flux with the charging and discharging capacities.

In the aim to include the **weather conditions** the global energy produced per hour and per technology is calculated multiplying the capacity installed by the hourly capacity factor. This factor is in fact the value of the production of one device at one hour. The closer to 1 the factor is, the better the weather conditions are i.e. the higher the amount of power generated is.

2.3 First results analysis

2.3.1 Composition of the energetic mix

In this section the first results of the model simulation are presented. An interactive function is used to view graphics. The following graph shows the electric mix of a week in November with RTE demand and the weather of 2006.

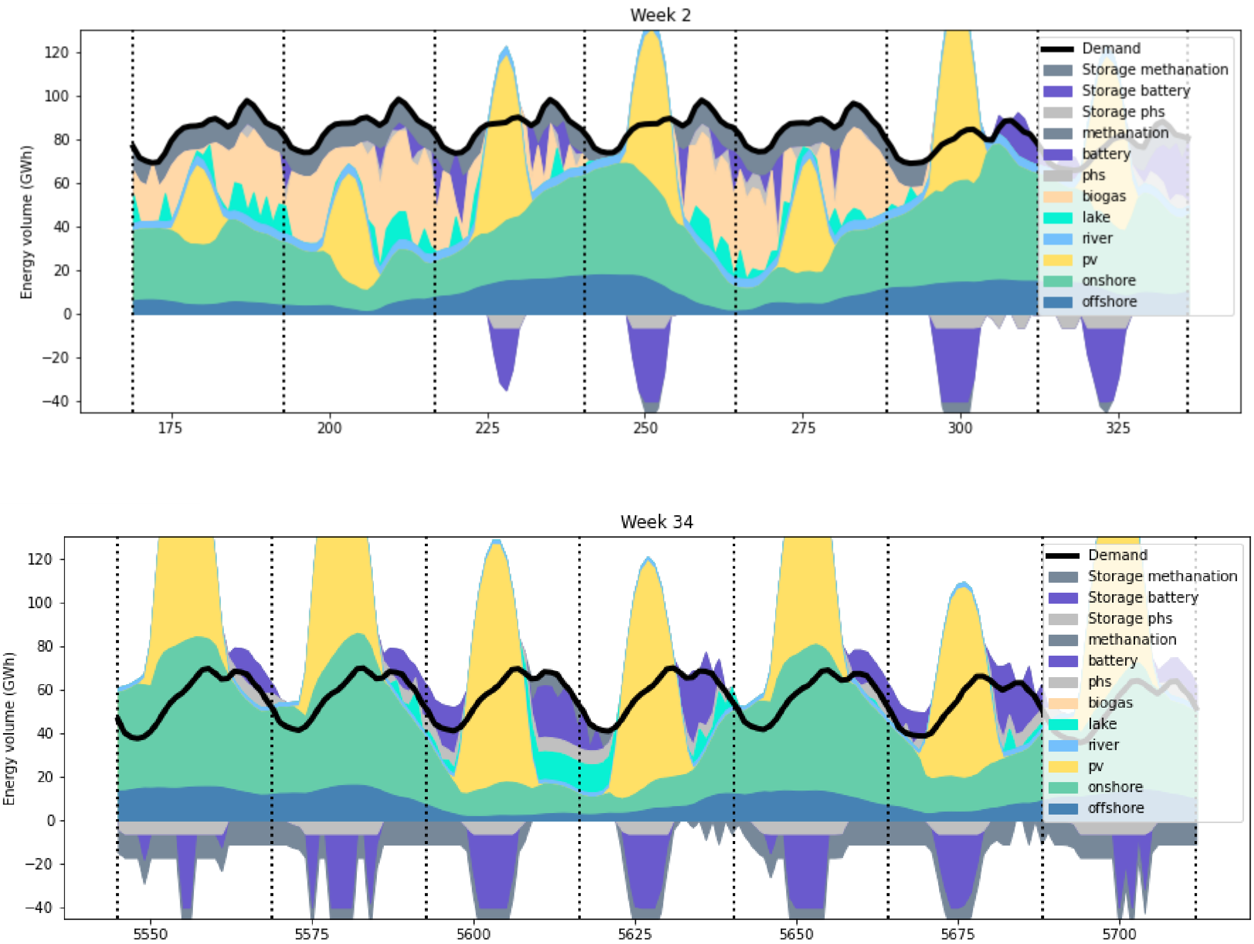


Figure 1: Both winter and summer energy mixes

These graphs represent the variations in energy input from each technology over a week. Firstly, the peaks in solar production are shown in yellow during the day; the reciprocal peaks in storage are shown, representing the excess solar production stored directly in the batteries. When there is a lot of wind then there is no need to destock to meet the demand.

The week 2 corresponds to the second week of January and the 34 corresponds to a week during August.

These weeks represent several phenomena:

- On days with both solar and wind power, we take the advantage of the weather to store in methanation, PHS and batteries.
- On days with very little offshore power, little onshore and solar we have to de-store a lot of energy and this mainly from lakes and rivers : long term storage.
- During winter days there is less solar power, to compensate it we use mainly biogas energy. Run-of-river energy is also used during winter time.
- During average days we store solar power in excess in battery and a bit in PHS and we use this storage during the night : short term storage.

2.3.2 Cost sensibility

Two main analysis can be done since the complete model and the model which only optimises the variable costs. For the complete model we look at the variability of initial capacity necessary to respond to the power demand for the whole year. The smallest model uses optimal values for this initialisation data therefore we compare the global cost after the optimization.

In the article more than 300 scenarios are created to evaluate the influence of the cost of each technology because of its uncertainty. To make this analysis I compared the influence of the price of one technology at a time with the central scenario.

Each technology has its own range of price variability.

Central scenario:

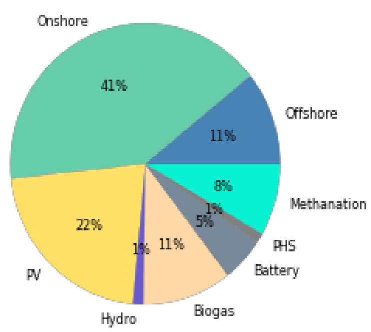


Figure 2: Central scenario

- Offshore :

Even if the price of offshore wind decreases, it will not be enough to make the mix differ significantly from the reference mix, this can be explain because of it reaches its maximum value allowed. If offshore wind has a higher price then it will disappear from the energy mix because it is too expensive.

- Onshore :

Even if we increase the price of onshore wind, it still has a large share in the mix; solar and offshore help to compensate for its decrease. As with the decrease in the price of solar panels, offshore wind disappears and is replaced by the technology whose price

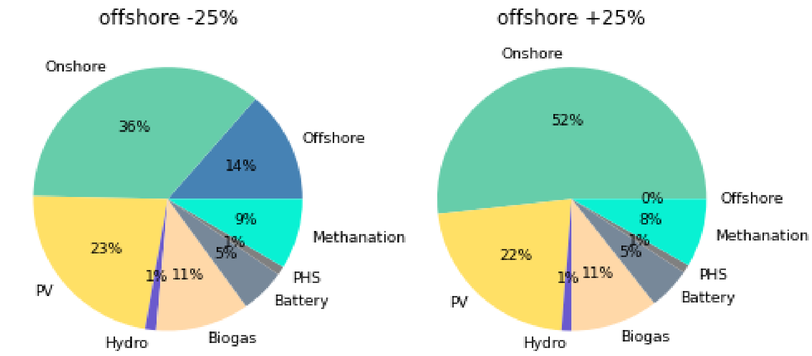


Figure 3: Offshore cost scenario

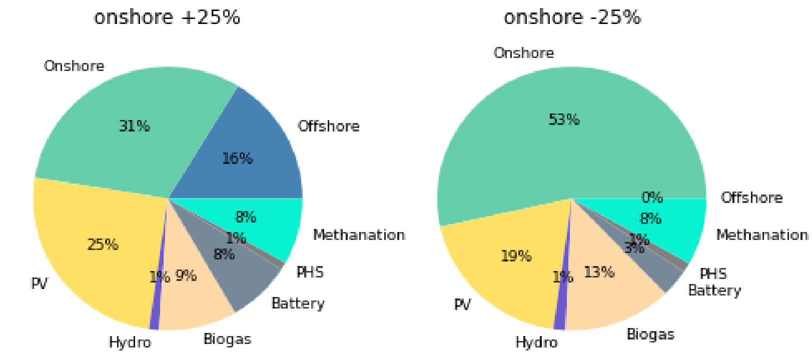


Figure 4: Onshore cost scenario

has decreased: onshore wind. The difference when solar has a lower share in the mix is the lower need for batteries which is compensated by biogas.

- Solar :

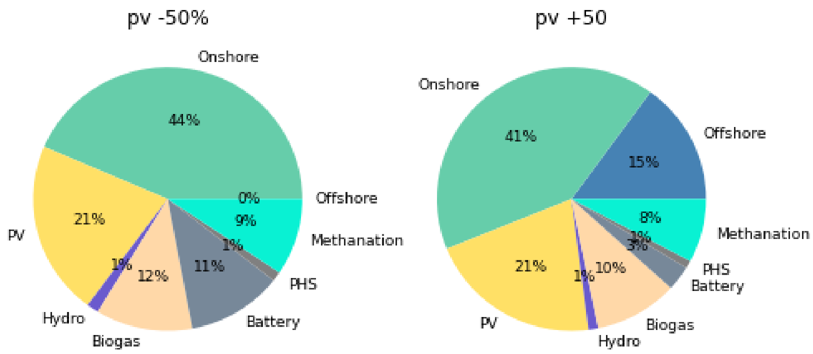


Figure 5: Solar cost scenario

The decrease in the price of solar panels leads to a strong increase in battery storage which allows for short-term storage of excess solar production; there is also a slight increase in biogas to compensate for the loss of offshore. The extreme increase in the price of solar panels does not change the mix, except for a slight increase in offshore wind.

- Battery :

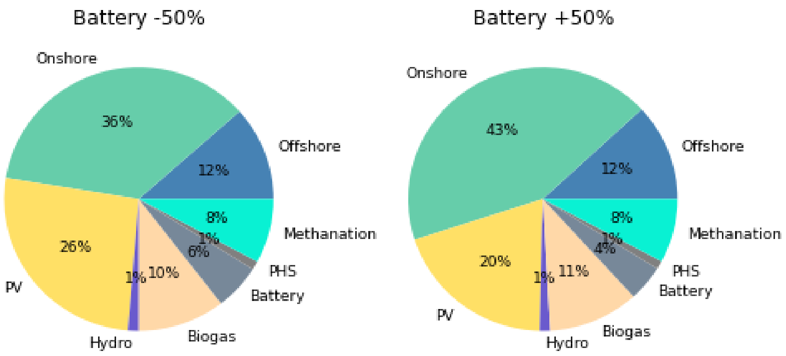


Figure 6: Battery cost scenario

Even with a sharp increase in the price of batteries its share in the mix remains indispensable. It can be seen that solar requires more battery storage than onshore wind, which is why the share of onshore wind increases when the price of batteries increases.

- Methanation :

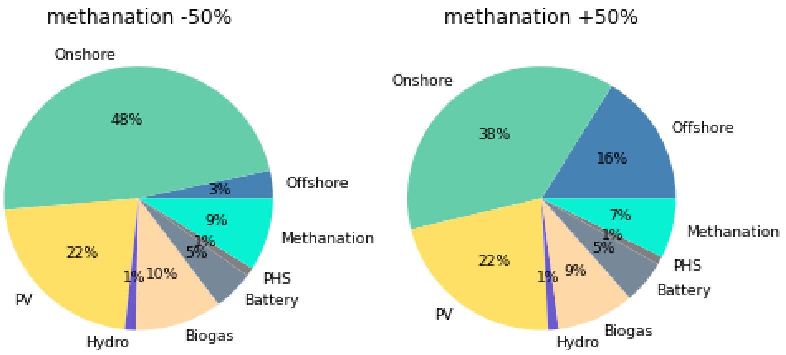


Figure 7: Methanation cost scenario

With a lower cost of methanation, the offshore technology decreases to leave more to the onshore because it is less expensive but with a greater need for storage. With a higher cost, the opposite phenomenon is observed.

2.3.3 Weather sensibility

Concerning the sensitivity of the model to weather, an analysis has been made in the EOLES model report. The graph on page 20 [1] shows the variation of the installed capacities according to the chosen weather year. It can be seen that the share of wind power in the mix is quite sensitive to weather variations; solar power is more stable (probably due to the regularity of its profile over the years).

Another analysis in the notebook Consumption_analysis of my Github shows that Negawatt's 2050 scenario is more sensitive to weather changes. This can be explained by the expected low demand for electricity which makes the mix more flexible without changing its price.

3 Production and consumption profiles

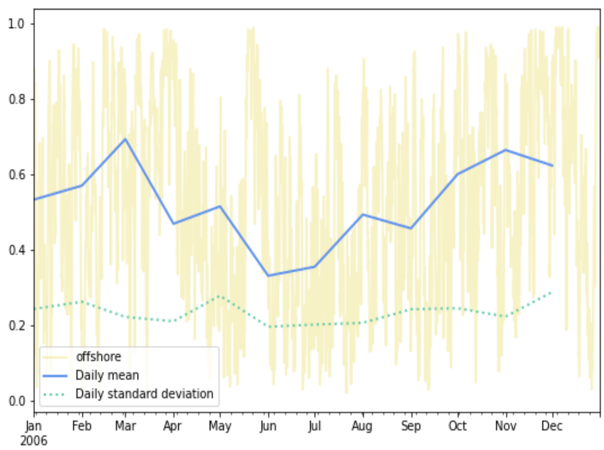
After an initial familiarisation with the model, I turned to the main objective of my internship: the modelling of new solar and wind production profiles. The aim was to test the robustness of the model trained so far with only 18 years.

3.1 Solar and wind power production profiles

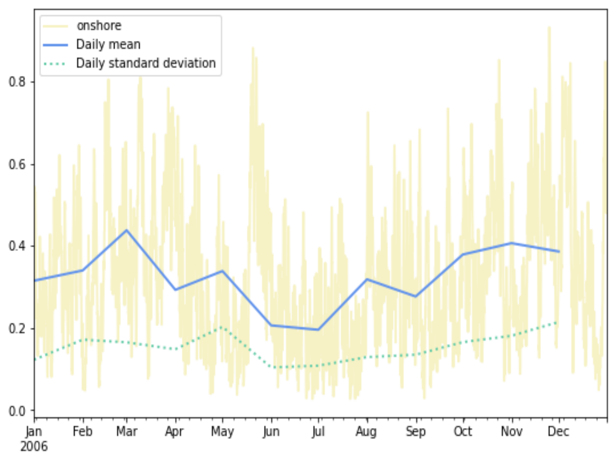
3.1.1 Profiles overview

The graphs show 3 curves: the raw weather curve (hourly capacity factor between 0 and 1), the monthly average of this weather and its standard deviation for solar and wind power of 2006.

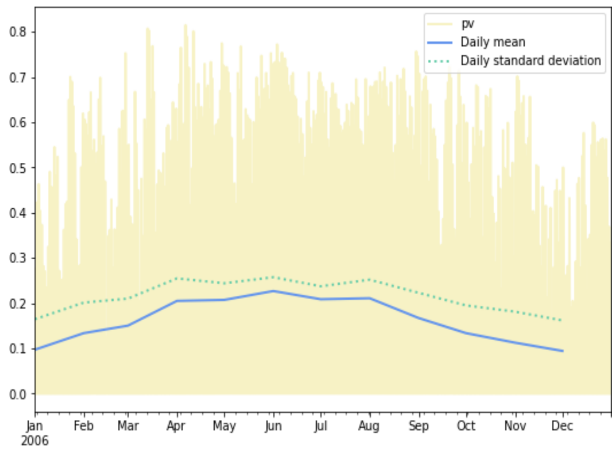
On the y-axis, we find the coefficient that qualifies the production at a given time h. This factor is then multiplied by the installed capacity to give the production in h. On the x-axis we find the 8760 hours of a year simplified by the monthly scale.



(a) Offshore production profile



(b) Onshore production profile



(c) Solar production profile

Figure 8: Production profiles

The graph of offshore production in 2006 shows a profile with high variability. There are months of the year with more wind overall and a fairly stable monthly variance. Over the year 2006, the capacity factor is on average 0.53 : the highest one among solar and wind power technologies, his yearly standard deviation is about 0.26.

The onshore profile is very close to the offshore in terms of trends (monthly mean and standard deviation), but on average its factor is less efficient than the offshore: 0.32 for the mean and 0.17 for the standard deviation.

The solar profile has a particular pattern because of its intermittency due to the alternation of day and night. Annual trends can also be distinguished, which will be discussed in more detail in the following sections. The capacity factor is the least efficient, probably due to the lack of production during the night: 0.16 for the annual average and 0.22 for its standard deviation.

3.1.2 Daily seasonality of renewable energies

To better understand the pattern of the weather I made a daily profile on several years (figure 9): it is calculated by averaging each hour of the day over a year. Among 7 years we can see a strong daily seasonality even for wind power.

For offshore there is a decrease of production between midnight and 10 a.m and then the capacity factor value goes back up until 20 p.m. Onshore daily seasonality is also significant. These variations can be explained by the correlation with air pressure and temperature. However, we observed differences between years, mostly for offshore power. For example, 2015 is a year with much more wind than 2017.

In the case of solar seasonality the pattern is very pronounced because of the alternation of day and night and there is very small differences between years. This seasonality will help to produce new profiles.

3.1.3 Annual trend

The following graphs (figure 10) show the annual trend of each technology over seven years. This trend is calculated by averaging the electricity production for one week.

From January to mid-May, the capacity factor of wind energy decreases; in summer it remains at low values and rises again in autumn.

For solar energy, the opposite phenomenon is observed, the curve is in the shape of a semicircle with a peak in summer.

3.2 Electricity demand for 2050

The graphs of electricity demand according to the scenarios in 2050 show the different more or less optimistic estimates. Here are the electricity demand figures:

- RTE: mean hourly electricity demand 66 GWh
- Ademe: mean hourly electricity demand 48 GWh

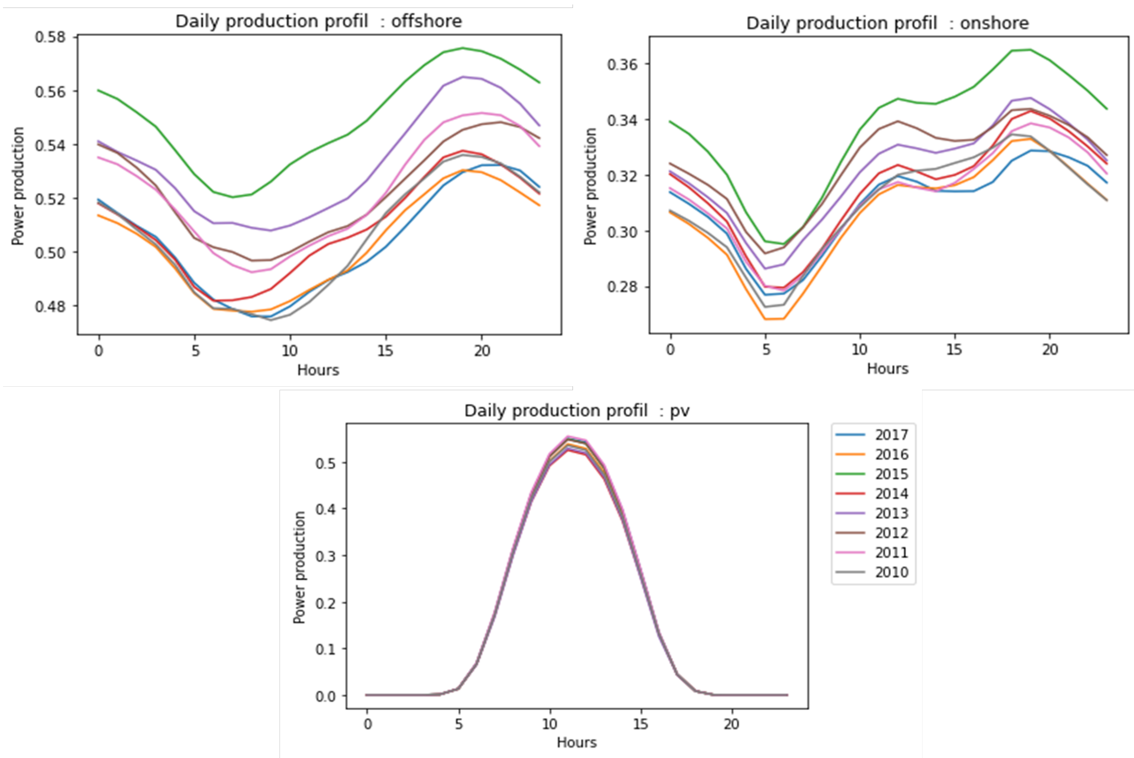


Figure 9: Daily power production profiles

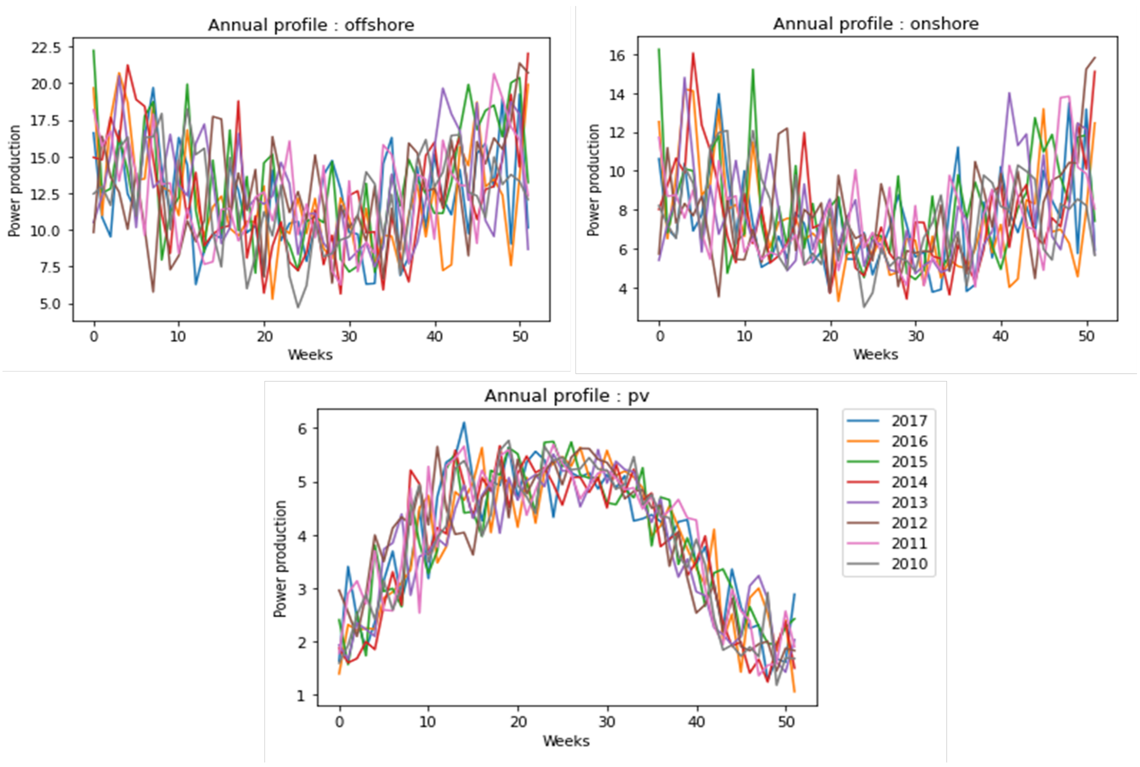
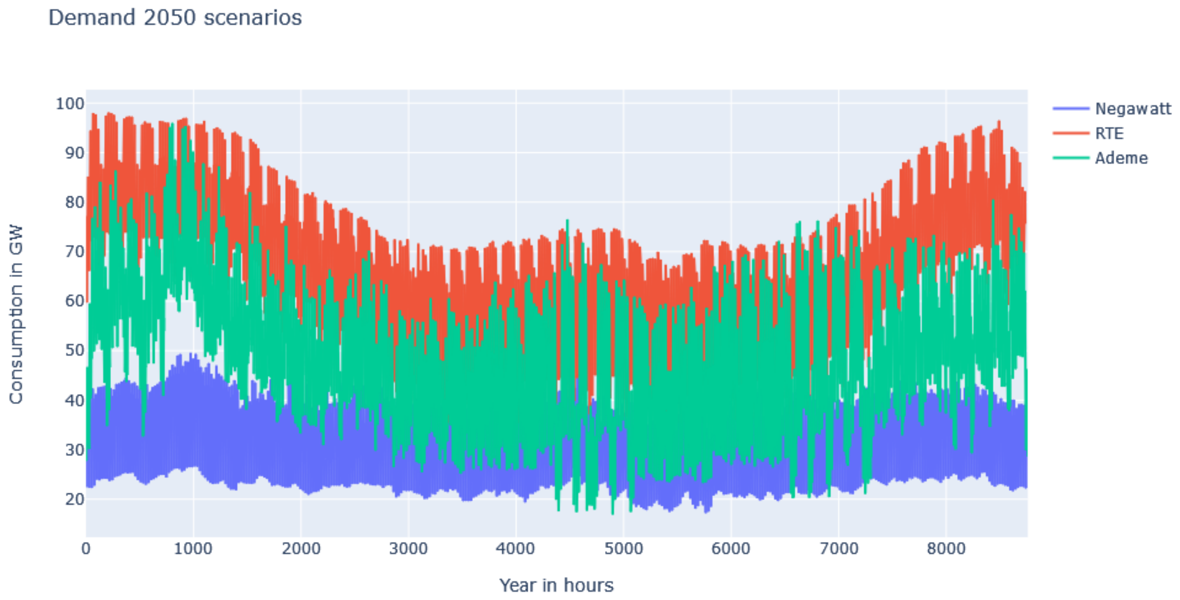


Figure 10: Annual power production profiles

- négaWatt: mean hourly electricity demand 31 GWh

We also notice the high demand in winter due to the use of heating systems, it should be noted that this difference could change given the increasing use of air conditioning in summer. This difference is very slightly observed for négaWatt, this is explained by the strong sobriety foreseen by the scenario: very little variation in the electrical demand.

If we zoom in on the profiles (second graph in figure 11), we can see the change in demand during the week for RTE and Ademe: less electricity is needed at the week-end (companies are not working, less daily travel, ...). This change is not observed for négaWatt because it makes the assumption that the electricity needs will have to be spread evenly over the week (working also on weekends).



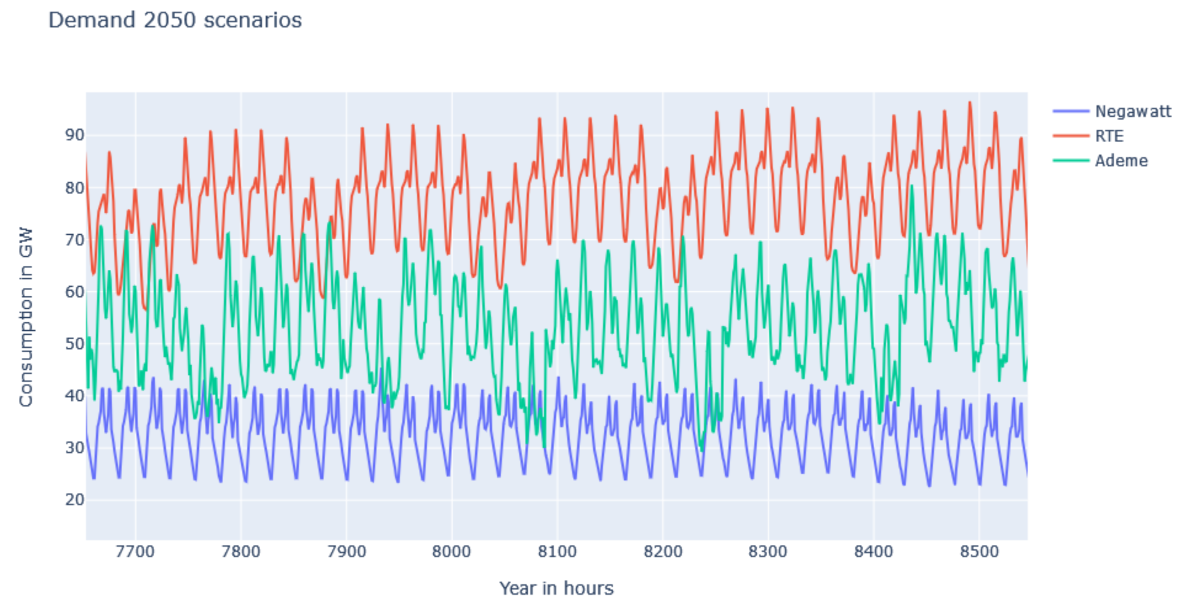


Figure 11

4 Production model

I will present the different methods studied to simulate new production profiles to test the robustness of the model in chronological order.

4.1 Modelling new profiles with Fourier transform

The first tool used to simulate new profiles is the Fourier transform. In fact, if we take a look at solar profile it is similar to a signal characterised by a frequency. The idea was to extract the main Fourier coefficients of the signal to reconstruct a new signal with discrete inverse Fourier transform. Here are the different steps of the method:

- New Fourier coefficients

First of all, to select the Fourier coefficients that will allow us to simulate a new profile, we retrieved the 8760 Fourier coefficients of 18 years. We use Python package *numpy.fft* which makes a discrete Fourier transform :

$$S(k) = \sum_{n=0}^{N-1} s(n) e^{-2i\pi k \frac{n}{N}} \tag{1}$$

with N = 8760, s the solar profile of one year.

From the 18 values of each coefficient we calculate the mean and standard deviation. Thanks to this information we can simulate new coefficients with a normal distribution. At this step we have a vector of 8760 new coefficients.

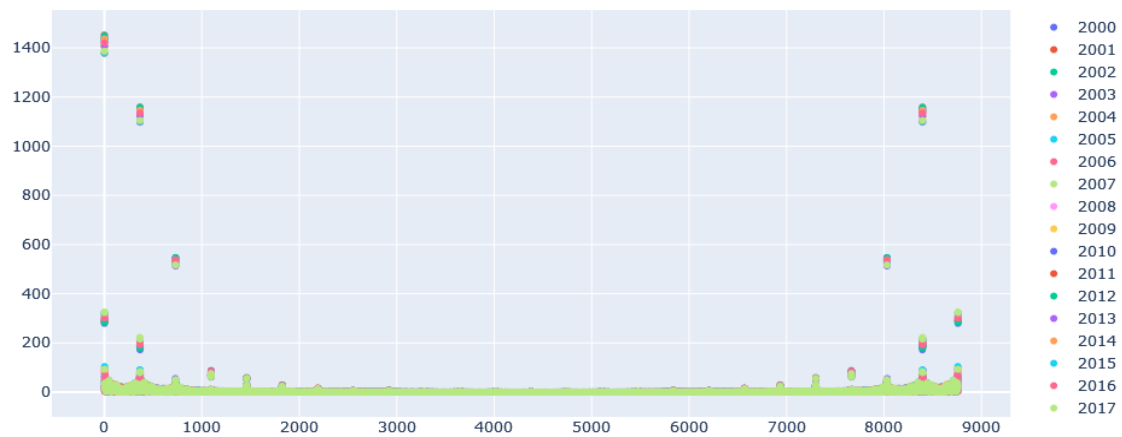


Figure 12: 18 years Fourier coefficients

This graphic show the module of the coefficients of 18 years. We can note that there is a vertical symmetry explained by the fact that the signal is real.

<u>Remark:</u> Since the basic signal corresponding to the solar production profile is real, we can use the Hermitian symmetry property :

$$S(k)^* = S(-k) \tag{2}$$

In fact, the other half of the coefficients (8760/2 until the end) are the complex conjugates from the first coefficients except the first one.

Indeed, we only have to simulate 8760/2 coefficients and to calculate the other half taking the opposite sign of the imaginary part.

- Discrete inverse Fourier transform

To reconstruct our signal we have to apply the discrete inverse Fourier transform on the new coefficients.

$$s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) e^{2i\pi n \frac{k}{N}} \tag{3}$$

- Results

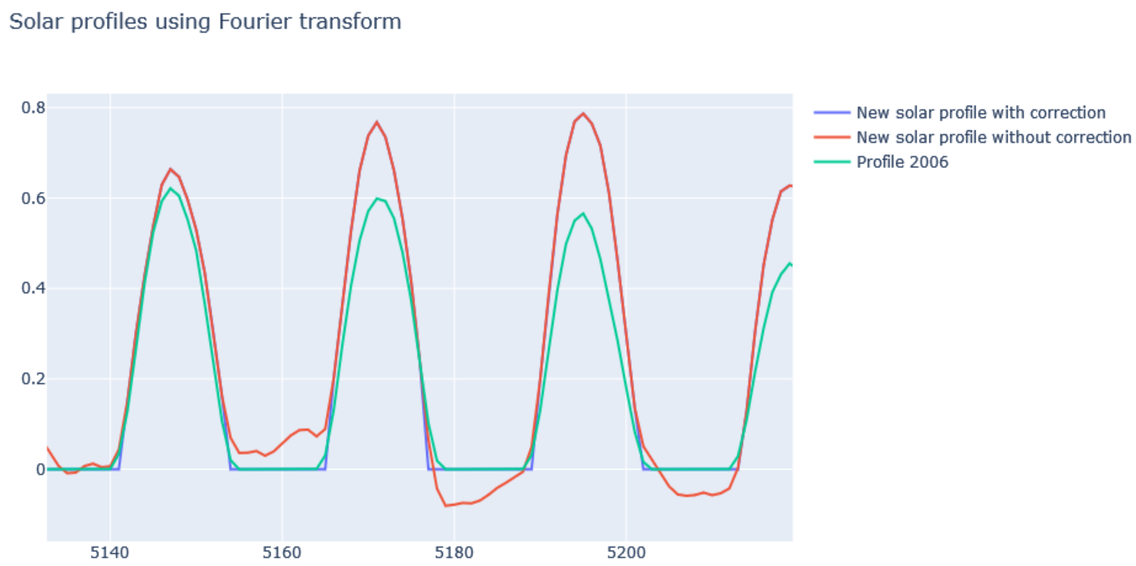


Figure 13: Zoom on new solar profiles

We can see that the new year of production is not always close to the profile in green which represents the year 2006. This difference is deliberate because we want to create a new profile.

<u>Remark:</u> Because of the regularity induced by the reconstruction with inverse Fourier transform we can observe negative values or positive ones during night time which isn't realistic. To correct this effect we force the values under a certain threshold to be zero.

On this graphic we can see both new and 2006 profiles with little variations but with annual and daily trends which seems respected.

- New profile analysis

The new profile produced using the Fourier transform is satisfactory in the sense that it reproduces a year of solar production while simulating a new weather hazard. The

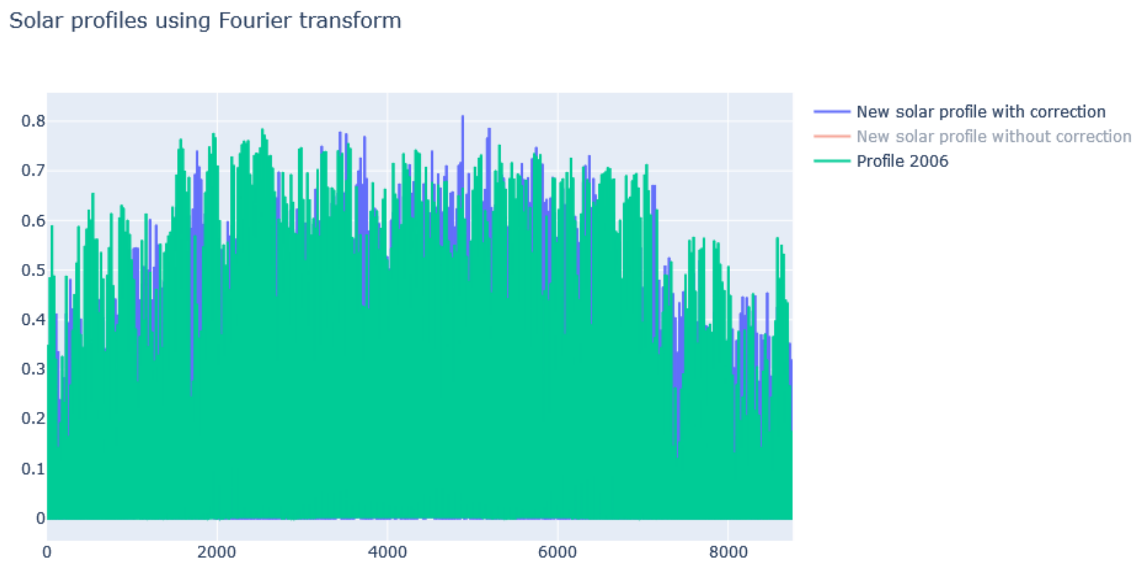


Figure 14: Solar profile simulated with Fourier transform

graphs in the notebook also show that the annual trend is respected and blends in with the other existing years. The day-night production pattern is also preserved thanks to the correction made to the raw reconstruction with the Fourier coefficients. The threshold chosen could be adjusted less coarsely. This empirical correction could be improved by forcing the model to simulate only positive values.

- Limitations for wind power

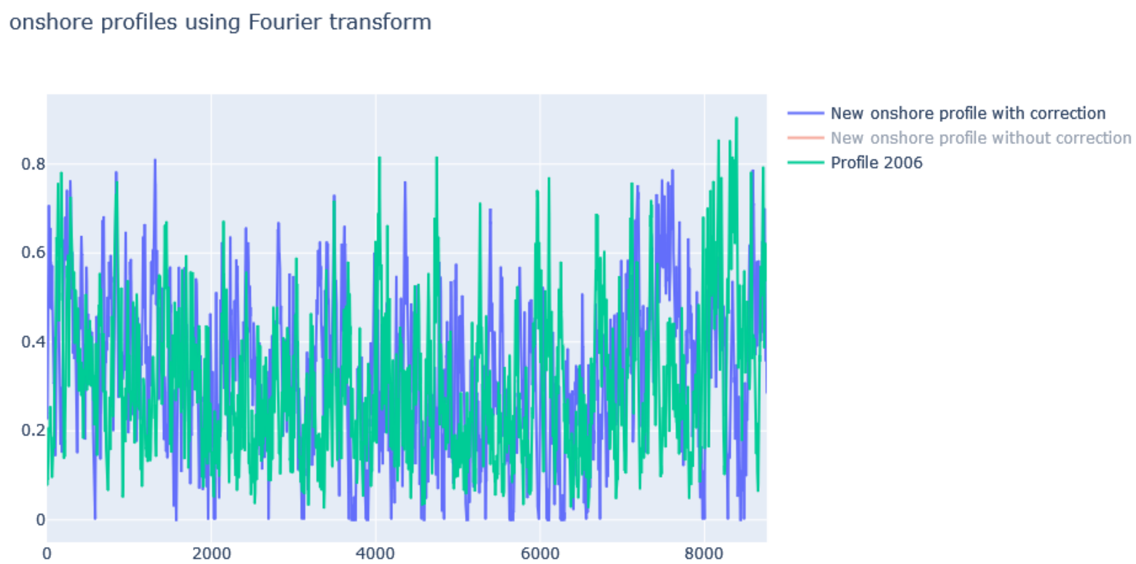


Figure 15: Onshore profile simulated with Fourier transform

Concerning wind power, the simulation of the new capacity factor is a little less satisfactory because applying the correction: constraining the values between the observed minimum and maximum (necessary to respect that it is between 0 and 1) implies constant values of wind power production which are not observed in reality. It is also observed that the profile is more dilated vertically. The Fourier coefficients method is more adapted to solar energy because of its sinusoidal structure and its simplicity to correct the night production peaks.

4.2 Modelling new profiles with Markov chain

After some bibliographical research on how to model wind power production profiles I found several articles explaining the interest of Markov chains. In particular, one article explains how by using second order Markov chains and adapting them to our dataset we can obtain realistic profiles [7].

A Markov process has the Markov property: the information useful for predicting the future is contained in the present state of the process and is not dependent on previous states.

A second order Markov process means that the future n+1 depends on the present n and the previous time n-1. The new approach consists in choosing a lag further than the previous one (non consecutive) which will allow to describe a more realistic future.

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, ..., X_{n-1} = i_{n-1}, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_{n-L} = i_{n-L}, X_n = i_n) \tag{4}$$

Simulation method:

- Data preparation

The Markov process, discrete in our case, is done on the residual of the time series (our production profile). To obtain it, different methods from time series are used to calculate a trend and a seasonality and to subtract them from the series. It is then necessary to check whether the residual constitutes stationary noise (i.e. whether its statistical properties (expectation, variance, auto-correlation) vary over time and whether their value is finite) on which we can apply the Markov algorithm.

- Data preparation

To get the states of the markov chain, one technique is to make the histogram of the values of a year and to choose the number of bins that will constitute the size of the state space.

Then we have to transform our dataset (18 years profiles) into chains with the states of one year.

- Build the transition matrix

The transition matrix of the chain is created from the following formula :

$$p_{ijk} = \frac{m_{ijk}}{\sum_{j=1}^{n} \sum_{k=1}^{n} m_{ijk}} \tag{5}$$

This formula correspond to the number of transition from i to j knowing that we were in i_{n-L} at lag L.

- Markov chain algorithm

With this matrix we create the cumulative transition matrix. We then initialise the new profile with L values. The algorithm consists in choosing a number between 0 and 1 through an uniform distribution, comparing this value to the values of the corresponding line (states) of the matrix.

- Limits

I tried to code this method but I faced several difficulties. I did not obtain any results using this method because, to obtain a profile relatively similar to reality, it is necessary to choose a large number of states for the Markov chain. A large dimension of the set of states makes the transition matrix of the chain grow considerably. The calculation of the coefficients requires to look for our 18 years of hourly production for each hour if a certain transition has already been observed. The construction of the matrix is therefore very time consuming. Moreover, it is possible that very few transitions determined in a single year take place in our data set, which allows us to construct the matrix. This lack of data implies many zeros in the matrix which makes the algorithm difficult to accept even when using a healing mechanism for unknown transitions.

The reasons for the failure of the method are also the discovery of a method more adapted to the use of the new production models in the framework of the EOLES model: typical days.

4.3 Modelling new profiles with typical days

The article [8] gave us the idea of another method adapted to the use of new profiles in the ENR project: simulating new weather by selecting typical days from the profiles in our dataset.

The principle is to select typical days based on the average and variance of daily production by clustering. To simulate new weather, new annual trend curves are simulated.

4.3.1 Modelling yearly trend

The first step is therefore to simulate new trends over the year. Since the clustering of the selection of days will be based on 2 criteria: the average and the variance of daily production, we need to simulate two new trend curves corresponding to these two criteria.

To simulate new profiles I decompose the time series into several signals: trend, seasonality and residual. The objective is to obtain a stationary noise, i.e. its expectation, variance and auto-correlation do not vary in time. The *seasonal_decompose* function of *statsmodels* [9] allows us to do this decomposition directly by forcing the noise to have this property. The following graph shows the different signals decomposed from the onshore weather over 18 years.

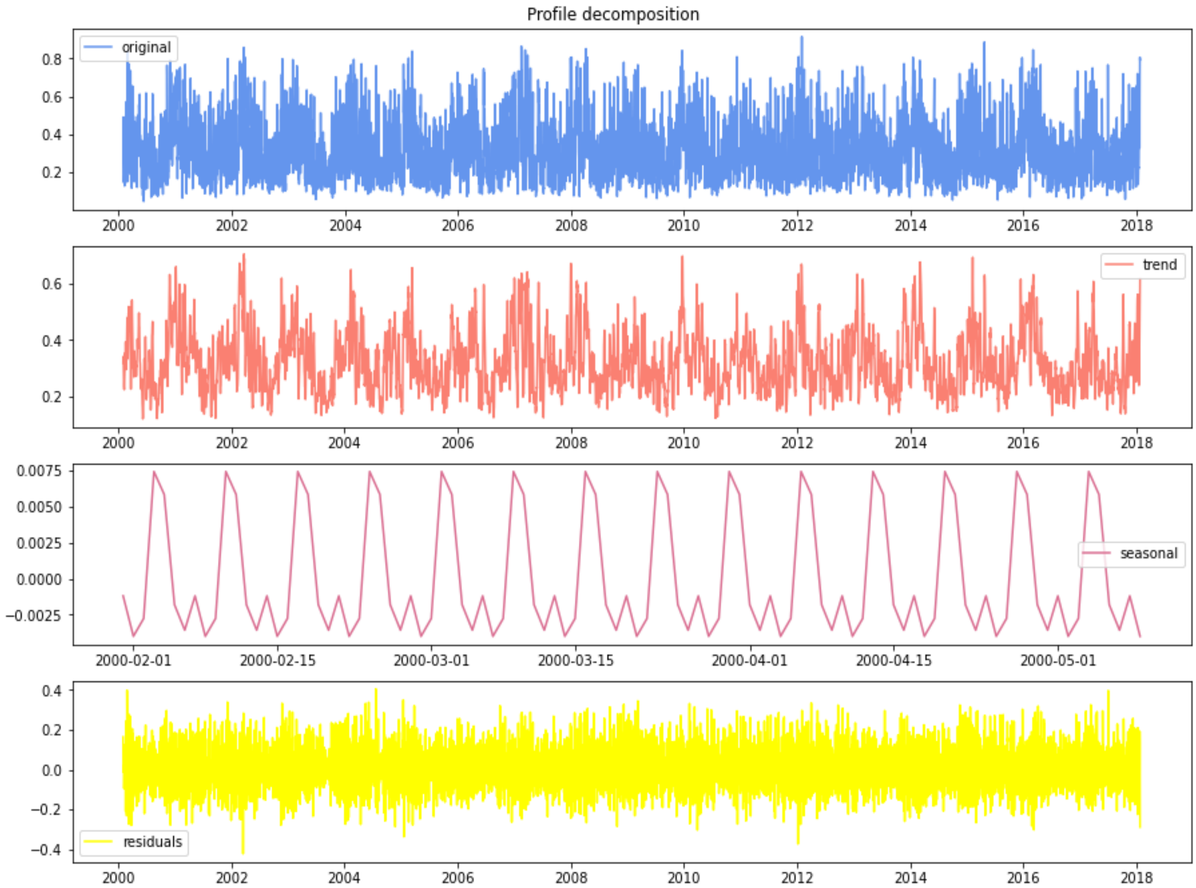


Figure 16: Onshore profile decomposition

From the white noise we recover its mean and standard deviation and simulate a new residual. We then re-multiply the trend and seasonality.

4.3.2 Typical days

One of the references in the article studied for typical days recommends the use of KMedoids clustering [10]. KMedoids package is available in the library *scikit learn* and the default metric used is Euclidean, I will come back in the next section to the interest or not of changing the metric. The simulations made in the article show that the adapted number of clusters is 12, however tests on this parameter for our dataset will be presented in the following.

The method used is an initial clustering on annual trends (daily production means and variances) over 18 years. The clustering returns centres associated with an existing day among the 18 years. We can thus find the hourly daily profile that corresponds to the typical day.

Once we have obtained our 12 typical days, we use the predict method of KMedoids: it allows us, from the clusters and centres already created over our 18 years, to associate a class to our 365 new simulated days with two new mean and standard deviation curves.

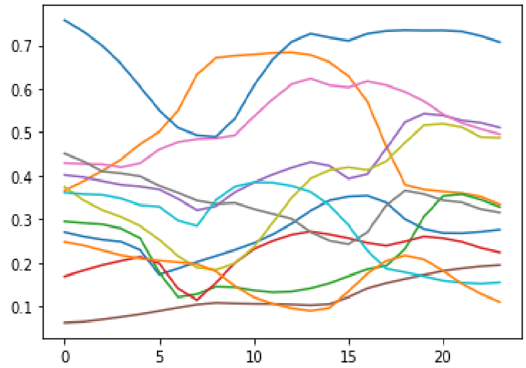
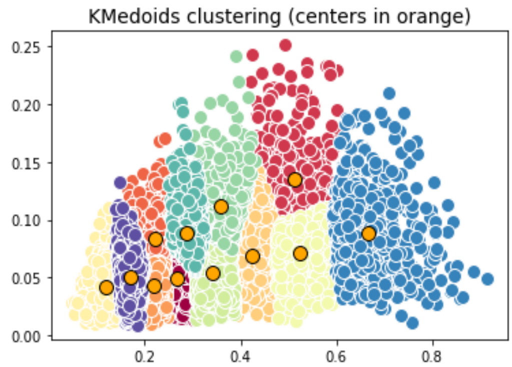
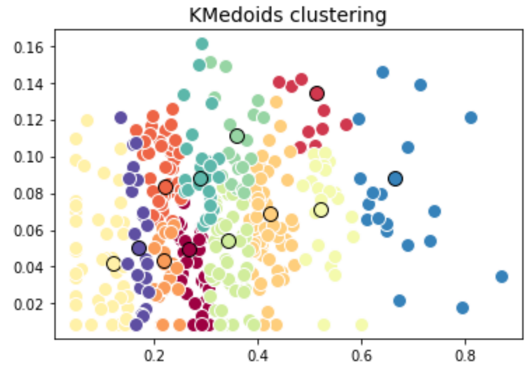


Figure 17: Onshore typical days



(a) Clustering with 18 years



(b) Clustering with the new trend

Figure 18: Clustering for onshore profile

The last step consists in associating each new day (characterised by its mean and standard deviation) with its associated typical day, i.e. the centre of its class.

The Figure 19 shows an other production year of onshore power. We can see another year to compare. The pattern seems to be very close from the one of 2006 with a different annual trend. Some tests were applied on the new profile: equality of standard variation, means and confirms the similarity of the profiles.

4.3.3 Parameters test on solar profiles

Parameters are involved in the clustering method:

- Metric definition

The number of typical days chosen is 12: the article "EnergyScope TD: A novel open-source model for regional energy systems" [10] explaining their use in the case of a model for strategic energy planning recommends a number of 12 typical days. I then carried out tests in the framework of the EOLES model which confirm that this number remains appropriate in our case.

- Number of clusters

The default metric used in KMedoids clustering is Euclidean. It could make sense to favour close production averages over a day compared to the daily variance. To do so, I

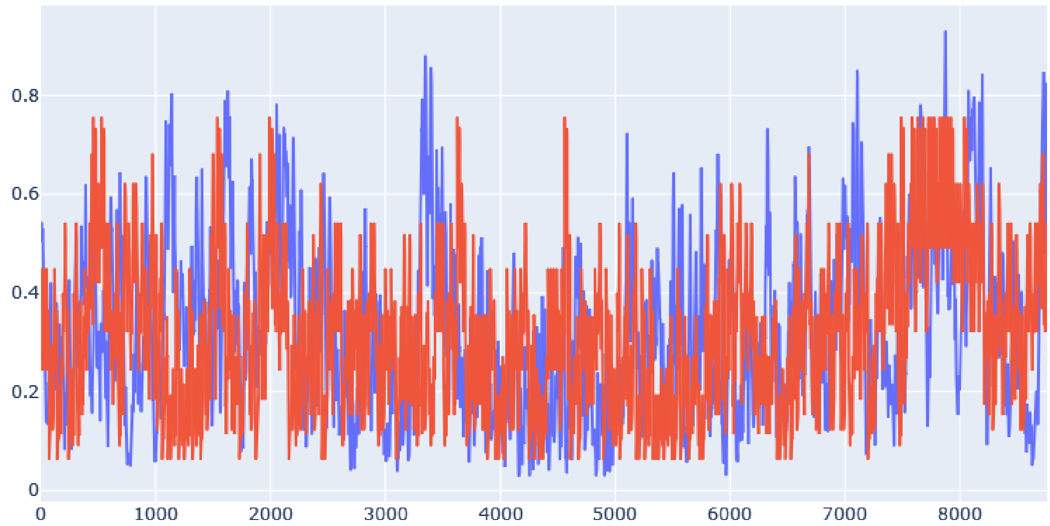
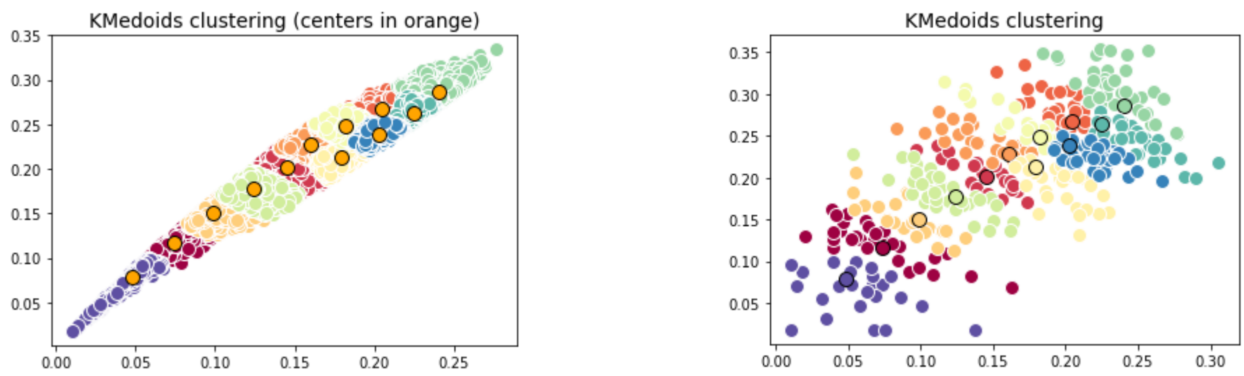


Figure 19: New onshore production year (red: typical year / blue: 2006)

modified the Euclidean metric by adding an alpha parameter in front of the variance data and giving it less weight. If we compare the difference between the reconstructed profile of an existing year without annual trend modelling we find that the error is greater if we give more weight to the mean if we choose 12 clusters. In the case of 6 and 18 clusters the value of alpha=1.5 allows us to obtain a lower error.

- Centers definition

The clustering of the typical solar days shows that the variance and the average daily production are strongly correlated. This correlation has been taken into account in the model because the centres of the selected clusters correspond to observations, which allows to recover the correlation lost when simulating a new annual trend of daily average production.



(a) Clustering with 18 years

(b) Clustering with the new trend

Figure 20: Clustering for solar profile

The centres are chosen by the clustering method but could be chosen manually. This model does not take into account extreme days since the centres will not correspond to outliers.

Moreover, it could also be interesting to choose days by looking at solar and wind productions simultaneously to reduce the size of the data: typical days of global weather (windless and sunny days, cloudy and windy, ...).

4.3.4 Continuity issue of wind power profiles

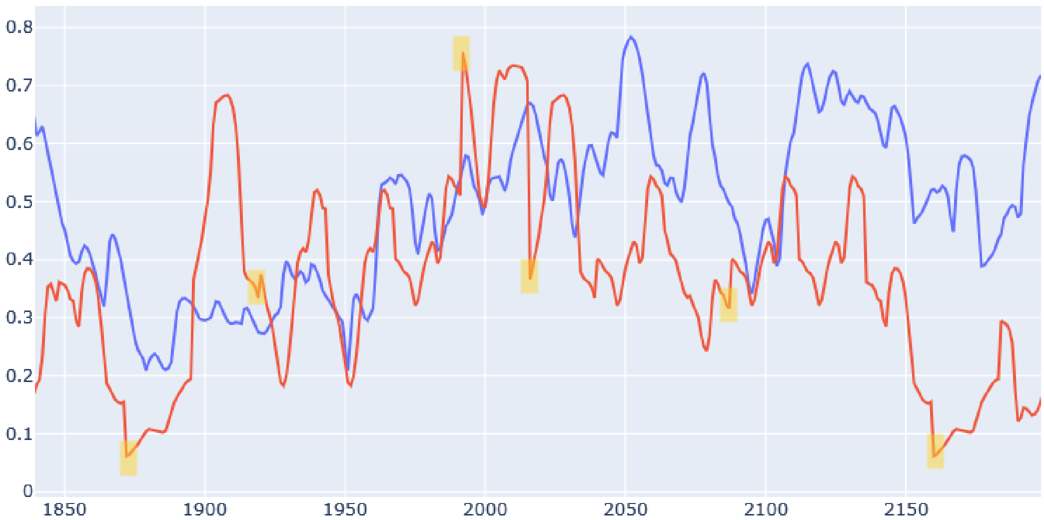
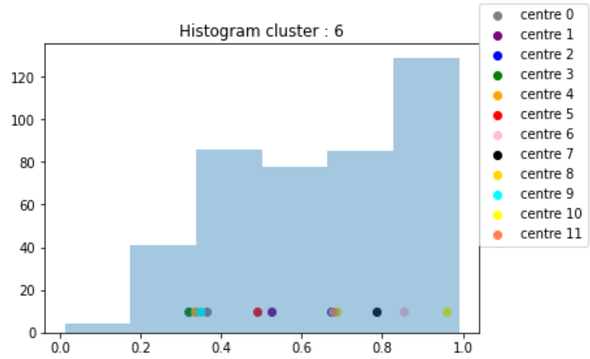


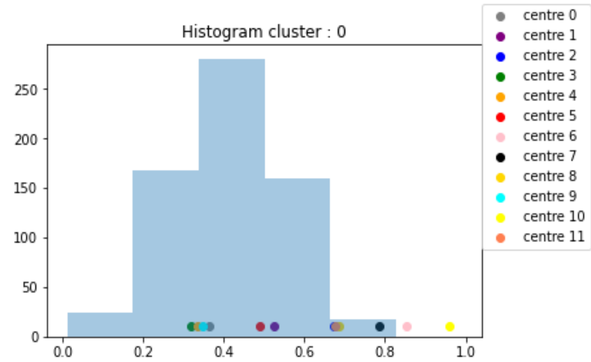
Figure 21: Zoom on onshore production year (red: typical year / blue: 2006)

On the wind profile simulated with typical days we observe discontinuities (highlighted in yellow on the figure 21) to smooth the profile we wanted to choose the centres of the classes of our clusters so that the recollection of the typical days is more acceptable: less difference in production between 11 pm and midnight.

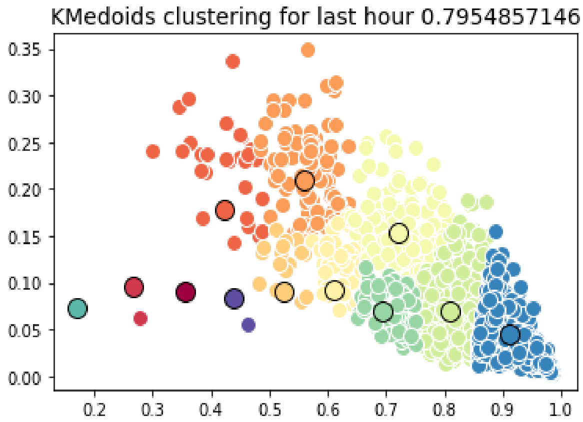
This modification in the clustering algorithm was added once the clusters were created by crossing the production groups at 11pm with the clusters resulting from the mean / variance clustering. It turned out that in some groups there were no observations.



(a) Histogram of offshore power production at 11 p.m. for 18 years of the cluster 6



(b) Histogram of offshore power production at 11 p.m. for 18 years of the cluster 0



(c) Clustering of offshore power production at 11 p.m for 18 years

Figure 22: Continuity issue of wind power profiles

On the 12 histograms obtained I have displayed the centres of the clustering of the means and variances. Thus, if we take for example a day in cluster 6, we associate its centre (in pink) which is located in the group with the highest production at 23h; by imagining that the following day is in cluster 0 (in term of means and variances of production), the centre of cluster 0 not being located in the group of highest production at 23h we will not be able to find a correspondence. Perhaps by changing the way the clustering is done or the associated metric we could find more acceptable centres for the recollection of days.

5 New simulations with typical days

To analyse the new production profiles I compared the cost by setting the installed capacity at the beginning of the year and then the cost and capacity by optimising both variables.

5.1 Cost variation

We compare the simulation of the optimal model (by fixing the installed capacities) by changing only the solar and wind production profiles here are the graphs of the cost variation.

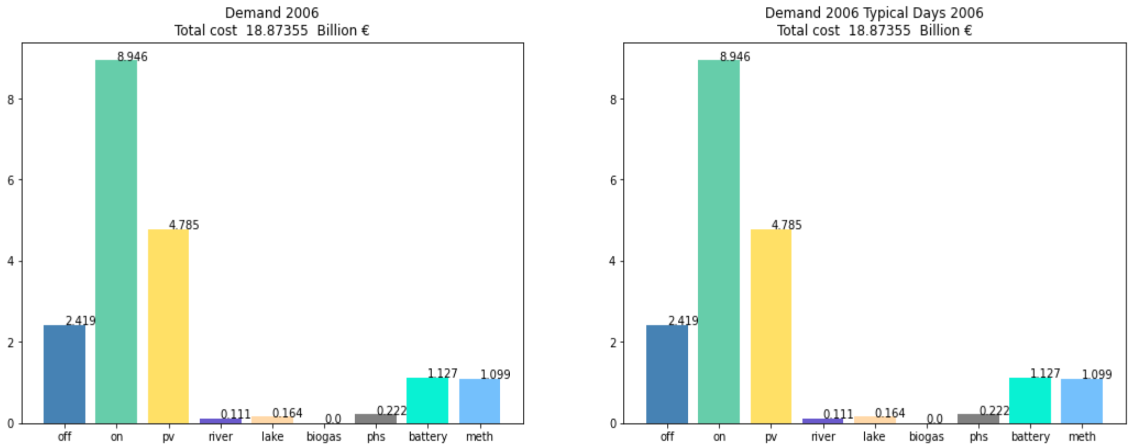


Figure 23: Cost variations

The typical days profile corresponds to the method described above but without the modelling of a new annual trend: the 2006 trend is used. It can be seen that the cost is not modified by the simplification of the profile into typical days.

If we now simulate 30 different weather conditions for each technology and compare the price obtained (we fix the installed capacity at the beginning of the year, it is the same for each simulation). The figure 24 is the boxplot of the mix cost obtained. We notice that the costs are between 28 and 29.2 billions of euros with an average at 28.3 billions euros.

5.2 Installed capacities variations

Now we simulate the complete model (cost and installed capacity optimisation). It can be seen that the simulation of a new production year with Fourier has more influence on the optimums than the use of typical days.

Figure 24: Cost variations for 30 new production years

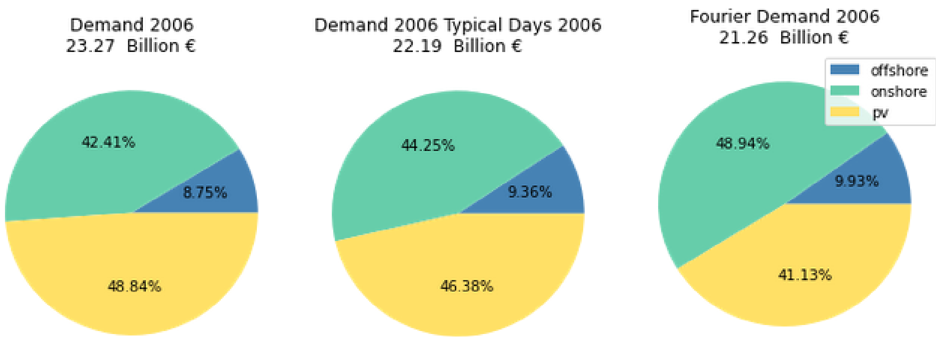


Figure 25: Capacities variations

The reason for the difference in mix is that when modelling a weather year with Fourier we do not rely on an observed year to model the trend, we simulate the coefficients using information from the 18 observed years.

If we look at the variation in cost for 12 randomly simulated weather years with typical days we see that the cost can vary between 26 and 28 trillion euros. The average is 27.5 and most of the costs vary with +/- 0.5 billion euros.

Regarding installed capacity, it can be seen that offshore does not vary: it reaches its maximum for each weather, onshore varies very little and solar has a wider range of values. These variations are compensated by the other technologies not represented on the graph in Figure 26.

Figure 26: Variations of the optimised variables (complete model)

5.3 Mix composition variation

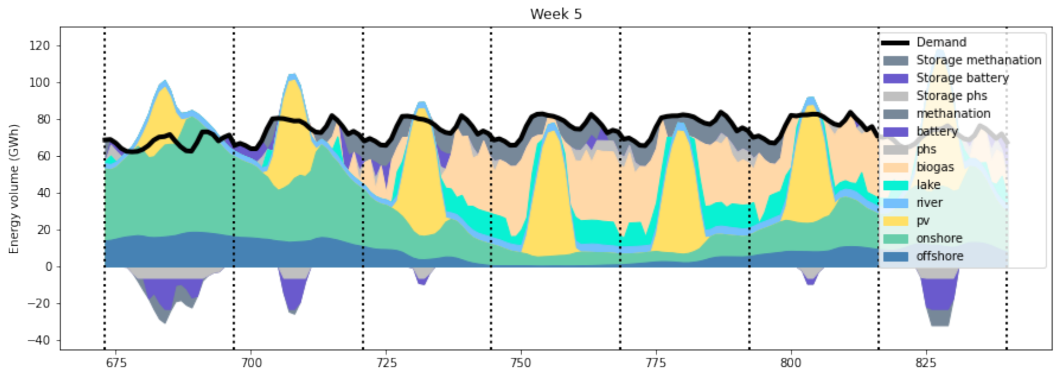


Figure 27: Mix with demand 2006 with production 2006

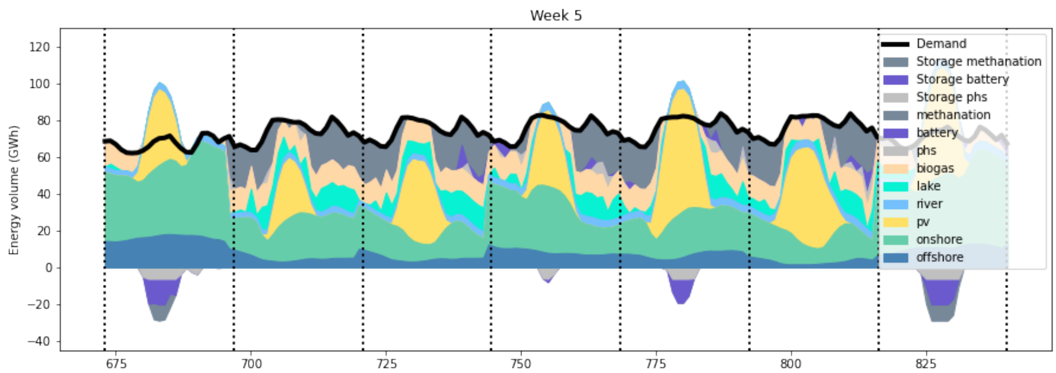


Figure 28: Mix with demand 2006 with new profile with typical days with trend of 2006

Here is the comparison of energy mixes for 2006 with actual data and those using typical days without new trend modelling. We notice the discontinuity of the wind profile, the composition of the mix varies slightly due to the difference in weather but the technologies used remain the same (in different quantities).

6 Limitations

6.1 Model limits

It could be interesting to optimise the mix in real time: in fact, in real life, the weather forecasts are known in the short term. Optimising the model by assuming that we know the hourly production over a year could distort the results. Taking this limitation into account would imply finding a new method to simulate weather over a few days only.

6.2 Modelling new profiles limits

The method for modelling trends has the advantage of keeping the correlation of the solar and wind profiles. Indeed, one year of production is chosen to simulate the 3 energies by simulating a new residual for each time series. However, the new trends remain similar to the existing years. It would therefore be interesting to find a better way of extracting the trend component of each production curve, i.e. with less information in this trend to actually simulate a new year.

Furthermore, it was noted that simplifying our profiles by using only 12 typical days per technology does not affect the cost and energy mix optimisation. It might be interesting to reduce the data size of the hourly profiles to daily averages by including the daily seasonality a posteriori.

Another questionable point of the new weather year modelling methods is that they do not explicitly include the correlation constraint between the offshore, onshore and solar capacity factors. The following table shows the correlation values between the day-averaged production curves for solar and wind power.

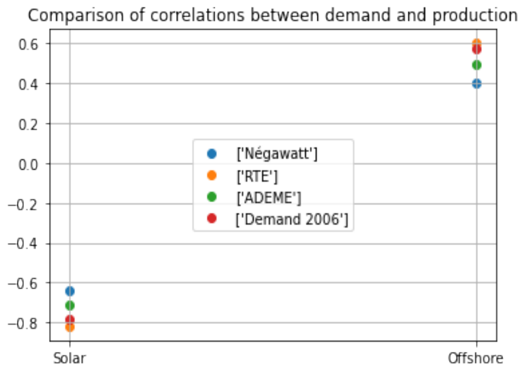
Tec	Onshore	Offshore	Solar
Onshore	1	0.91	-0.38
Offshore	0.91	1	-0.40
Solar	-0.38	-0.40	1

Table 1: Correlation of solar and wind daily power production

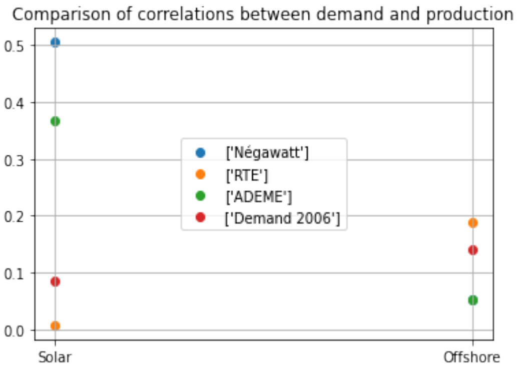
It can be seen that the correlation between offshore and onshore production is very strong and that the correlation between solar and wind power exists and cannot be neglected. They are even stronger if we look at the correlation of the time profiles. Taking this correlation into account therefore seems to make sense.

Tec	Onshore	Offshore	Solar
Onshore	1	0.99	-0.65
Offshore	0.91	1	-0.74
Solar	-0.65	-0.74	1

Table 2: Correlation of solar and wind monthly power production



(a) Correlation of monthly mean profiles



(b) Correlation of hourly mean profiles

Figure 29: Clustering for onshore profile

If we take a look at the correlation between electricity supply and demand, we can also see that they are linked. Very strongly on the one hand, and this for any scenario, because the demand for electricity depends on the seasons (monthly profiles): we heat more in winter. On the other hand, the correlation on the hourly profiles varies according to the chosen scenario, so the simulation of these new profiles have more or less influence depending on the strategy used (which is consistent with the observation made for négaWatt in part 5).

Conclusion

To conclude, this laboratory internship allowed me to approach mathematics applied to environmental issues. I was able to mobilise my knowledge in optimisation and statistics and my coding skills. My work as an intern, which was based on the EOLES model, had the objective of simulating new weather years, i.e. capacity factors for onshore and offshore wind power and for solar power.

To reproduce new time series I first analysed profiles for the production characteristics: annual and daily trends and seasonality. The extraction of these patterns allowed me to use different methods: the Fourier transform adapted to solar profiles with the advantage of not being influenced by the production trend of an observed year. Indeed, the second method that led to exploitable results is the typical day method. This technique requires the modelling of new annual trends, this step has to be improved.

The analysis of the EOLES model simulations with new weather profiles shows first of all that for a fixed mix the replacement of a year of observed production by the simplification in tpyical days has no influence on the cost of the mix: the same quantity is generated overall for each technology. If we look at the variations in installed capacity for 12 typical days, we see that among solar and wind, only the quantity of solar panels installed is sensitive to the weather. The cost optimum found also varies around 27 billion (with an average standard deviation of 0.5 billion).

The techniques used to model new weather patterns have their limits: the correlation between the profiles of each technology and electricity demand is not negligible. It also seems advisable to review the method of modelling the annual trend to reproduce years that are independent of those observed.

However, we have been able to show that in the framework of the EOLES model, hourly weather data could be simplified. It could be interesting to reduce the size of the weather data to optimise the calculation time.

References

[1] P. Q. B. Shirizadeh, Q. Perrier, "How sensitive are optimal fully renewable power systems to technology cost uncertainty?" *Energy Journal, International Association for Energy Economics*, no. 43, 2022. [Online]. Available: https://doi.org/10.5547/01956574.43.1.bshi

[2] "Production totale," 2020. [Online]. Available: https://bilan-electrique-2020.rte-france.com/production-production-totale

[3] Leréveilleur. (2021) Quelle électricité pour demain ? [Online]. Available: https://www.youtube.com/watch?v=XbrcUz0pu80

[4] "Scénario négawatt 2017-2050," *Institut Négawatt*, 2017. [Online]. Available: https://negawatt.org/IMG/pdf/synthese_scenario-negawatt_2017-2050.pdf

[5] "Transition(s) 2050. choisir maintenant. agir pour le climat," 2022. [Online]. Available: https://transitions2050.ademe.fr

[6] Hourly power output from wind and solar pv farms. [Online]. Available: https://www.renewables.ninja/

[7] H. A. T. Pesch, S. Schröders and J. Hake, "A new markov-chain-related statistical approach for modelling synthetic wind power time series," 2015. [Online]. Available: http://dx.doi.org/10.1088/1367-2630/17/5/055001

[8] H. J. Gauthier Limpens, Stefano Moret and F. Maréchal, "Energyscope td: A novel open-source model for regional energy systems," 2019. [Online]. Available: https://doi.org/10.1016/j.apenergy.2019.113729

[9] Seasonal-trend decomposition using loess (stl). [Online]. Available: https://www.statsmodels.org/dev/examples/notebooks/generated/stl_decomposition.html

[10] A. C.-A. F. Domínguez-Muñoz, J. Cejudo-López and M. Gallardo-Salazar, "Selection of typical demand days for chp optimization," 2011. [Online]. Available: https://doi.org/10.1016/j.enbuild.2011.07.024

Github link:

`https://github.com/AmandineCapitaine/Project_Renewables_2050.git`