ECOLE POLYTECHNIQUE
PROMOTION X 2021

MAP595 - Mathématiques financières
Daniel Sbeiti

RESEARCH INTERNSHIP REPORT

EDF Lab Asia Pacific - Singapore

Pricing of 24/7 Carbon Free Energy Contracts

Abstract

In the light of increasing government carbon regulations and green investment culture, companies like Google have set ambitious targets to achieve 24/7 Carbon Free Energy (CFE) within the next decade, necessitating precise and efficient matching of energy demand with renewable sources on an hourly basis. The Lab's mission is to address this demand-supply gap through its 24/7 CFE contract.

This study's primary objective is to develop a pricing model for these contracts, minimizing the investment costs required to ensure optimal management of renewable energy production and storage assets covering a contracted share of demand. The research focuses on resolving a non-linear stochastic optimization problem with storage control. Yet the complexity of the model requires developing numerical approaches. The implemented method is quite new in this field, as it combines deep learning neural network with optimal control.

Future work involves scaling the pricing algorithm to cover a full year, refining its complexity and precision, and potentially developing a fully operational model for commercial use by EDF. This project not only highlights the practical application of deep learning in market analysis and pricing, but also aligns with the evolving landscape of quantitative finance, where machine learning methods are increasingly employed to create innovative profit strategies.

Acknowledgments

I would like to deeply thank:

- Edouard Lavillonnière - *Director of EDF Lab Pacific Asia and internship tutor:*

  For his warm welcome to the lab and close supervision throughout the internship. As well as for his quality insights and guidance on both technical and managerial aspects. He taught me to confidently take the lead in a project and efficiently communicate progress with colleagues.

- Henry Zhu and Peixuan Wang - *Product Manager and Quantitative Researcher at the Lab, as well as team colleagues on the project:*

  For their support, help and guidance on this project. For welcoming me as part of the team and sharing with me the healthy spirit of a research lab. For their constant availability for help and advice providing managerial methods and technical insights.

- Xavier Warin - *Researcher at EDF R&D France:*

  For closely following my work throughout the project and providing me with expert advise on implementing the numerical methods. For constantly reaching out to help me fully grasp the methods of his research paper and helping me develop sharp skills in coding, deep learning and optimization.

Contents

1 Introduction

1.1 Introduction to EDF Lab R&D Singapore

The EDF Lab in Singapore is a Research and Development center operating in all Pacific Asia under the Electricité de France group (EDF). The group is globally known for its low carbon energy initiatives, and the Singapore Lab focuses primarily on developing sustainable energy technologies in Pacific Asia. The Lab often partners with top universities, experts or research institutions in common projects to accelerate the adoption of sustainable energy practices.

1.2 Market Context of 24/7 Carbon Free Energy (CFE)

Government regulations and policies

Governments across the globe are increasingly stating the importance of transitioning to CFE in a common drive against climate change. Policies and regulatory frameworks on companies' green supply coverage are therefore being implemented. On one hand, carbon pricing mechanisms like carbon taxes push companies to reduce their carbon emissions benefiting green energy resources. On another hand, government subsidies and tax credits encourage the development and adoption of green energy technologies. This context leads to the development of smart grids and energy storage to facilitate the adoption of renewable energies.

Raising commitments from corporations

As a consequence, major companies are making significant commitments to 24/7 CFE to meet raising sustainability goals and green investment expectations of stakeholders. For example, Google and Microsoft have announced 24/7 CFE by 2030, Amazon by 2025. These companies thus aim at covering all their energy demand with renewable sources on an hourly matching.

1.3 The Lab's mission:

The Lab's 27/7 project aims to bridge the gap between this described Renewable Energy demand from companies and an Renewable Energy supply. This service operates as a "24-7 Carbon Free Energy" (CFE) **contract** sold by EDF, and signed for a year between:

- One or several consumers who wish to cover a share $\rho \in [0, 1]$ of their live energy demand with Renewable Energies on an hourly matching. We will call ρ the CFE score. Consumers can be mainly corporations seeking to meet sustainability goals, especially in the light of the increasing regulations on sustainability reporting and sustainable investment.
- A 24-7 service provider who contracts with consumers to meet their Renewable Energy demand. The provider must match or exceed the contracted demand, while minimizing its supply costs, by optimally managing Renewable Energy (RE) loads from:
  - owned RE production assets or contracting with RE producers.
  - owned energy storage assets or contracting with energy storage operators.

The objective of the internship is to price these contracts for EDF to be able to commercialize it to its clients. The contract price of \$/MWh is deduced from marginal pricing, based on the optimal investment cost level. The optimized investment cost level should allow enough green resources to cover the agreed share of demand (the *CFE score*). These investment costs account for both energy production plants (solar panels, wind farms...) and batteries acquisition (we will neglect operational costs.

2 Problem formulation and Modelling

2.1 The model

We first decided to model the market through three key processes:

- An electrical demand $(D_t)_{t\in[0,T]}$
- A green production $(R_t)_{t\in[0,T]}$
- A battery storage profile $(u_t)_{t\in[0,T]}$

Each battery has an injection/withdrawal rate of 1MW per hour, for a capacity of n hours. u describes the injection ($u > 0$) or withdrawal ($u < 0$) of electricity in the batteries at each time step.

Whereas $(D_t)_t$ ix an exogenous process that the producer can't control to achieve the contract, the producer can work on the production and storage in order to achieve an agreed coverage of the demand in green energy. Let's further define two variables:

- Q the average level of production on a period T
- $\bar{X}$ the number of batteries we are working with

The producer's investment cost would depend directly on the choice of Q and $\bar{X}$, which hold a double condition: they must be high enough to allow enough production and storage to satisfy the green demand, but not too high as it is directly impacting the producer's cost. Since we define the price of the contract as the investment cost for the producer to satisfy it, We must determine the optimal cost of satisfying the green demand. We therefore optimize on Q and $\bar{X}$, which is an optimization problem under constraint:

$$Cost = \inf_{Q,\bar{X}} f(Q,\bar{X}) \tag{1}$$

$$E\big[g(\rho,(D_t)_t)\big] \leq \sup_{u} E\big[h((R_t)_t,(u_t)_t,(D_t)_t)\big] \tag{2}$$

$$\forall t \in [0,T] \quad u_t \in [U_t^I, U_t^W] \tag{3}$$

$$\forall t \in [0,T] \quad dX_t = -u_t dt, \;\; X_t \in [0, n\bar{X}] \tag{4}$$

Remarks and model analysis

- Equation 1 is the cost function of the optimization problem. f is the investment cost to achieve an average production level of Q MW and operate $\bar{X}$ batteries.
- Equation 2 is the main constraint that the parameters have to respect. It is called the "green constraint". The left term represents the overall green demand, as a percentage ρ of the total demand D. The right term represents the green offer, combining production R and storage u.
  Function h also takes demand D as an argument, this accounts for the fact that at each time step t, excess green production after covering the step's demand can't be used to cover demand at future time steps. This *"non green washing condition"* makes the constraint function h non-linear which adds complexity to the problem.
- Equation 3 is the constraint on the control $(u_t)_t$ that we need to generate. Batteries have time dependant bounds on injection (I) and withdrawal (W). We have $U_t^{W,I} = U\big(R_t, D_t, \bar{X}, (u_k)_{k\in[0,t]}\big)$.
  - *Injection quantity* is limited by excess production, battery availability and hourly injection rate.
  - *Withdrawal quantity* is limited by unmet demand, stored energy availability and hourly withdrawal rate.
- Equation 4 is the dynamic of the battery level $(X_t)_t$ as an accumulation of injections given timely withdrawals. It will help assert battery storage or stored energy availability for injection and withdrawal.

2.2 The processes

To model demand and production, we chose to generate Ornstein-Uhlenbeck processes, with α an efficiency rate (usually 0.7) and Q the average level of production:

$$dR_t = \kappa_r(\alpha Q - R_t)dt + \sigma_r dW_t \tag{5}$$

$$d\tilde{D}_t = -\kappa \tilde{D}_t dt + \sigma_d dB_t \tag{6}$$

$$D_t = \tilde{D}_t + f_t^D \tag{7}$$

The parameters for production have been calibrated on wind farm production data. The demand process is calibrated on the French 2013 historical consumption data.

We generate the processes using Euler's formula (and $\Delta t = 1$):

$$R_{t+\Delta t} = R_t + \kappa_r(\alpha Q - R_t)\Delta t + \sigma_r\sqrt{\Delta t}Z_t^r \tag{8}$$
$$\tilde{D}_{t+\Delta t} = \tilde{D}_t - \kappa\tilde{D}_t\Delta t + \sigma_d\sqrt{\Delta t}Z_t^d \tag{9}$$
$$Z_t^r,\, Z_t^d \sim \mathcal{N}(0, dt = 1)$$

Here is what we obtain as green production and demand curves over a year, using a mean parameter $Q = 45000$ and a monthly reversion rate:

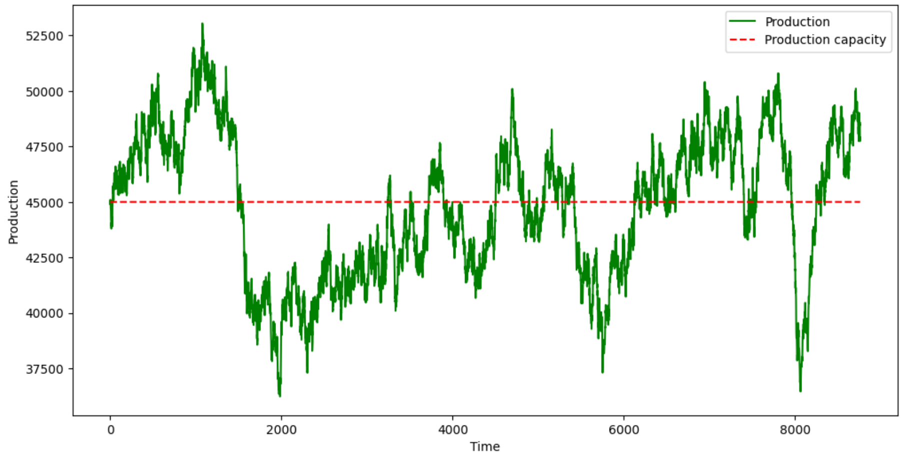

Figure 1: Green production curve over a year period

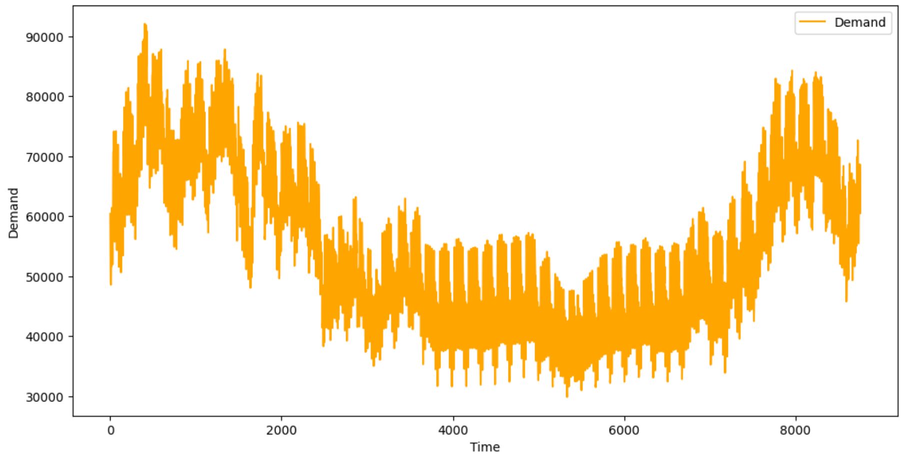

Figure 2: Electrical demand curve over a year period

Since out battery injection and withdrawal capacity is dependant on excess production or lack of production compared to demand, let's plot $(R_t - D_t)_t$ to analyse the battery relevancy in a general context:

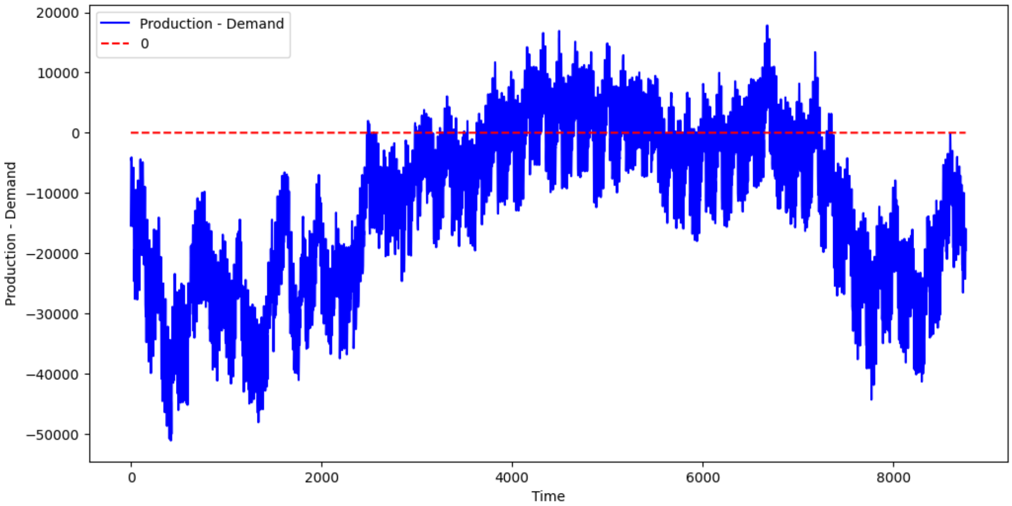

Figure 3: Excess/lacking production compared to demand on a year period

We can see on Figure 3 that the lacking production typically happens in winter where demand is very high (beginning and end of period) in France as people will use heaters in every closed place. Whereas summer period presents an excess production compared to demand. Battery storage would allow to store the summer excess production to help satisfy the higher winter demand and reach the agreed green demand coverage at lower investment costs, instead of just investing on higher production level calibrated on winter demand.

2.3 Previous and Related Work

Literature on contract pricing and market making in this format is quite poor. Pr. Nizar Touzi initiated this project and internship last year to work on the same problem **without** storage in batteries. After exploring some numerical resolution approaches, they ultimately achieved an analytical resolution of the optimization problem.

The consideration of a storage profile control to the optimization problem now adds a whole new level of complexity to it which makes it utterly way too technical to solve analytically. It is now more interesting and practical to explore numerical resolution methods.

Yet optimization problems with storage is a well known category of problem and is still very new and unexplored. The optimization cost function being non-linear, we can't use classical optimization algorithms like simplex based methods and have to explore new creative approaches. Literature on this problem category is very poor, yet Xavier Warin documented in documents [4] and [5] a numerical approach for reservoir optimization facing the market. His papers describes the profit strategy of storing energy in batteries when market price is low and selling when market price is higher. The resolution method is in that case an optimization problem under storage control very similar to ours, and Xavier Warin decides to solve it using stochastic gradient descent on neural networks. To fully grasp the described methods I gathered knowledge on stochastic gradient descent optimization algorithms in document [3]. It also happens that Xavier Warin is a quantitative researcher at EDF, we therefore had the chance of working closely under his supervision to adapt and implement his approach on our task.

The work realized in the precedent internship cannot be directly reused during this one as the model is now solely different. Yet it can be used as a tool for result analysis and comparisons for more insightful optimization of the algorithm. Combined with Xavier Warin's papers, these two research pieces constitute the pillars of my work on this project.

3 Numerical approach

3.1 Reformatting the model

In the numerical approach we implemented, $(u_t)_t$ is generated through neural networks. We decided to use one neural network per time step. We then use gradient descent optimizers to compute optimal Q, $\bar{X}$ and the NN parameters generating $(u_t)_t$.

This methods needs us to compute at each epoch a loss function that would account for both minimizing investment costs, and penalizing unsatisfied green demand. We will therefore use as loss function our cost function in equation 1, penalized with the green demand constraint in equation 2, by writing the Lagrangian format of the constrained optimization problem:

$$Cost = \min_{Q,\bar{X}} \left\{ f(Q,\bar{X}) - \max_{\lambda \geq 0} \left\{ \lambda \times \left(\sup_{u} E\big[s\big(\rho, (R_t)_t, (D_t)_t, (u_t)_t\big)\big] \right)^{-} \right\} \right\} \tag{10}$$

With $s = h - g$, for h and g from equation 2.

We only account the penalization inside the loss function if it yields a negative value, thus the max with 0. Indeed, we do not aim to reward the loss if the condition is met, we only aim to penalize if it is not met (to respect the direction of the inequality in 2).

The parameter λ

The λ parameter here measures the importance accorded to fulfilling the demand constraint in our cost optimization as a "duty cost".

The optimization problem operates in two steps. We first fix λ at a high value (not too high to avoid regularization, but not too low to keep importance on the constraint). Since f is independent from λ, performing a min-max inversion in equation 10 allows us to determine the $\min_{Q,\bar{X}}$. We then solve the final loss for different values of λ to determine the $\max_{\lambda \geq 0}$ and get the final cost.

3.2 Neural networks and optimizers

The control intervenes in non-linear expressions, making the optimal control problem very tricky to solve analytically like it was previously done in the problem without storage. We therefore propose to solve the problem numerically. The approach is the same as in the paper on storage management facing the energy market strategy described in document [4]:

We define T neural network models, one for each time step, that we train to deliver the time dependant control profiles $(u_t)_t$. At each epoch, the cost function in equation 10 is used as loss for updating the gradients to calculate Q, $\bar{X}$, and the parameters of the neural networks. Equation 10 is minimized through stochastic gradient descent. At the end of the training, we should obtain the optimal Q,$\bar{X}$ and control profile $(u_t)_t$ to achieve an empirical CFE score ρ equals to the set target.

Here is below the pseudo-code of the algorithm implemented to solve the first part of the optimization (i.e with fixed λ):

Algorithm 1 Stochastic Gradient Descent with ADAM for Optimal Control

Initialize: Variables initial values $q \leftarrow Q_0$, $\bar{x} \leftarrow \bar{X}_0$
Initialize: Neural networks $(\text{NN}_t)_{t=0}^T$ of parameters $(\theta_t)_{t=0}^T$
Initialize: ADAM optimizers $(\text{adam}^{\theta_t})_{t=0}^T$ for NN parameters
Initialize: ADAM optimizers adam^q, $\text{adam}^{\bar{x}}$ for Q and $\bar{X}$

for $i = 0, \ldots, M_{\text{epoch}}$ **do**
  $L \leftarrow f(\bar{x}) + g(q)$ (*initial loss*)
  $X \leftarrow 0$ (*initial reservoir*)
  $R \leftarrow \alpha q$, $D \leftarrow 0$ (*initial values*)

  for $t = 0, \ldots, T$ **do**

    Forecast production and demand:
    Generate Z_r, $Z_d \sim \mathcal{N}(0, 1)$
    $R \leftarrow q\big(\alpha + (R - \alpha)e^{-\kappa dt} + \sigma_r\sqrt{\frac{1-e^{-2\kappa dt}}{2\kappa}} Z_r\big)$
    $D \leftarrow De^{-\kappa dt} + \sigma_d\sqrt{\frac{1-e^{-2\kappa dt}}{2\kappa}} Z_d + f_t^D$

    Compute the storage controls:
    $\hat{u}_t \leftarrow$ Generate in $[0, 1]$ using $\text{NN}_t(R, D)$
    $U_t^W \leftarrow U(R, D, X, \bar{x}, q)$ (*energy stored* $\wedge$ *lacked demand* $\wedge$ *withdr. cap.*)
    $U_t^W \leftarrow U(R, D, X, \bar{x}, q)$ (*left storage* $\wedge$ *excess prod.* $\wedge$ *inject. cap.*)
    $u_t \leftarrow -U_t^I + (U_t^W + U_t^I)\hat{u}_t$ (*value now lies in* $[-U_t^I, U_t^W]$)

    This allows to capture bound conditions in eq. 3 according to doc. [4]

    Update battery storage state:
    $X \leftarrow X - u_t$

    Update loss:
    $L \leftarrow L - \lambda\Big(\delta g\big(\rho, R, D, X\big)\Big)^-$
  end for

  Update q **and** x **with gradients from** $\text{adam}^q(L)$, $\text{adam}^{\bar{x}}(L)$
  $q \leftarrow q - \eta_q^i \nabla_q^i L(q, \bar{x}, \theta)$
  $\bar{x} \leftarrow \bar{x} - \eta_{\bar{x}}^i \nabla_{\bar{x}}^i L(q, \bar{x}, \theta)$

  Update NN parameters $(\theta_t)_t$ **with gradients from** $(\text{adam}^{\theta_t}(L))_t$
  for $t = 0, \ldots, T$ **do** $\theta_t \leftarrow \theta_t - \eta_{\theta_t}^i \nabla_{\theta_t}^i L(q, \bar{x}, \theta)$ **end for**

end for
Output: $Q = q$, $\bar{X} = \bar{x}$, and $\{u_t\}_{t=0}^T$

3.3 Limits

While trying to implement this model, I stumbled on a couple limitations that I had to bypass in order to obtain a functioning algorithm.

Processes stochasticity:

Let's remind that the main goal here is trying to minimize the penalization of not reaching the objective green coverage percentage ρ in constraint equation 2, by updating the parameters Q through stochastic gradient descent. We then use this new Q to generate a new production process that we test for constraint 2 until condition met. When performing a stochastic gradient descent, the norm is usually to generate only one stochastic process variable per step to simulate the expected value we are working with.

Yet a single stochastic Ornstein-Uhlenbeck process R of mean level Q can generate a very high variance of empirical ρ values due to its stochasticity. These values are used to calculate the penalized constraint, and would sometimes deliver empirical ρs higher than the objective at early stages **(variance superior than distance to objective)**. This means that some steps randomly satisfied the constraint just because of the variance on R's simulation giving counter intuitive instructions to the optimization model which couldn't converge (cf below figure).

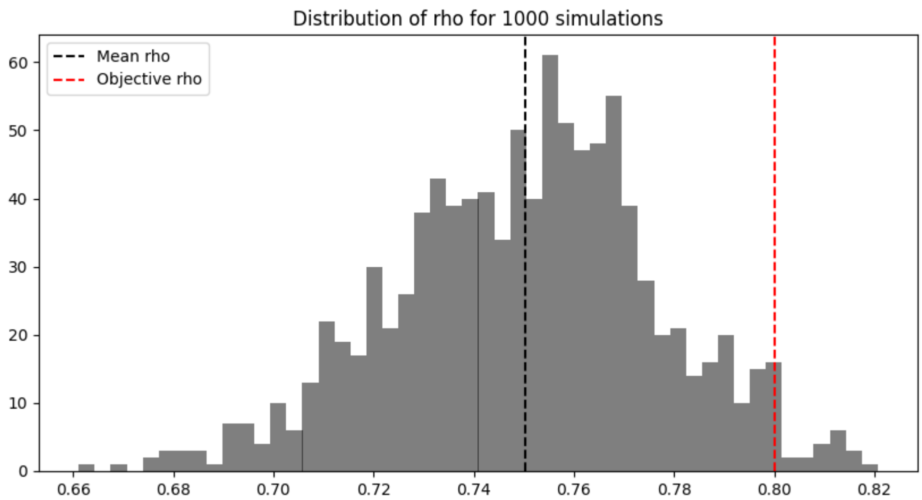

Figure 4

We therefore had to reduce this variance on empirical ρ in order to coherently compute the penalized constraint and make the model converge. A way to do this is to batch the code and simulate the expected value as an average of m simulations. Indeed we have that, for $X_1, \dots X_m$ i.i.d:

$$V\left[\frac{1}{m}\sum_{i=0}^{m} X_i\right] = \frac{1}{m}V\left[X_1\right] \tag{11}$$

We now obtain a calculation of empirical ρ with very high precision ($std = 0.008$ vs 0.14 before) using a batching of size $m = 300$, and can converge to an optimal solution precisely.

Lack of excess production:

As precised earlier, the battery relevancy is directly correlated with the presence of an excess production to be stored. If production is never in excess compared to the demand, then the battery is never charged and will never be used. Yet the production level relies on the average production parameter Q, optimized to satisfy a percentage ρ of total demand.

We have yet observed that if ρ is too low, constraint on production requires a low Q value which determines production level. Production is then too low to ever be in excess compared to demand at any time step and the battery is never used. In that case, we are back with the initial problem of optimization without storage dealt with in last year's internship.

To determine after what approximate ρ level batteries become irrelevant, let's plot the total excess production compared to demand as function of ρ. To obtain this curve, we first determine Q associated with each ρ using the model without storage solved analytically last year, and we calculate production using that mean parameter Q to obtain total average excess production over the given time period.

Here is the graph obtained:

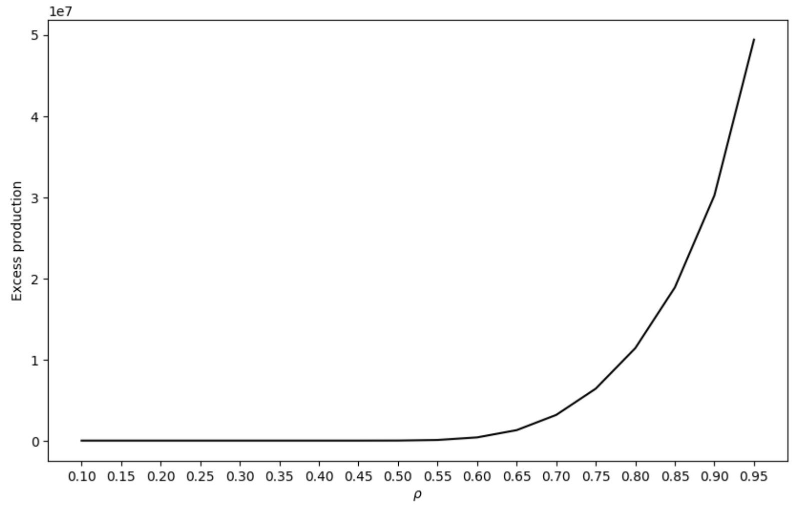

Figure 5: Excess production compared to demand with ρ

As observed on the above figure, excess production on the overall year period only starts to appear above ρ constraints of 0.55. When considering only the time range we are working with (50 time steps), the minimum ρ for batteries becomes 0.75.

Computational price:

One of the biggest bottlenecks to obtaining a good algorithm was the computational cost of training. We were calling our neural networks 70000 epochs $\times$ 300 batches $\times$ 8760 time steps. These add up to more than a hundred billion iterations for one run (i.e a couple days). We needed to get that duration down in order to fine-tune it.

A first solution is to first operate on a limited number of time steps like 50, and then use splitting methods to scale it up to the whole period.

Another technique used was to kickstart the algorithm with a high learning rate on early epochs allowing big initial steps in the optimization. Then slowly get the learning down as we progress to end up with a precise evaluation. The function used for the learning rate comes from document [2] and is called 'linear cosine decay', I discovered it watching the algorithm time optimization video [1].

Here is below the decayed learning rate used in the algorithm:

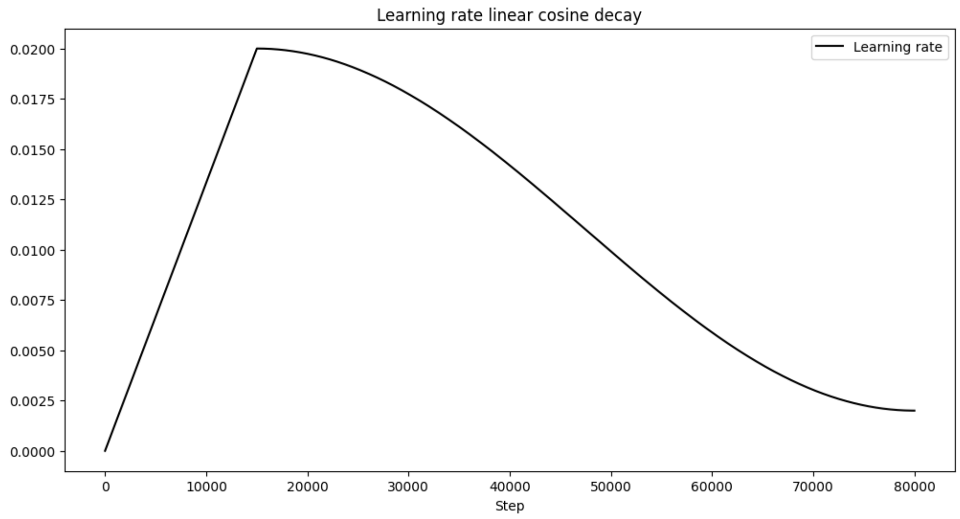

Figure 6: Linear Cosine Decay for optimizer learning rate

These among other techniques helped us get the code down to operating in around 2 hours allowing us to run it multiple times in a day and fine-tune it using results analysis.

4 Results

4.1 Presenting the optimized solutions

After adjusting all hyper-parameters, we obtained a converging algorithm after for 50 time steps that we will need to scale up to $t \in [0, T]$. In order to visualize its results, I decided to plot 9 relevant graphs:

Plots with regards to time to visualize the final behaviors of our system after optimization:

- Final demand $(D_t)_t$ and green offer to grasp that the green constraint is verified (covering the agreed CFE score)
- Excess production $(R_t - D_t)_t$ to analyse battery relevancy
- Reservoir level $(X_t)_t$ to visualize battery usage
- Optimized storage profile control $(u_t)_t$ to complement precedent analysis

Plots with regards to training epochs to visualize the evolution and convergence of our model:

- Evolution of our optimized variables Q and $\bar{X}$ to monitor their convergence and final values
- Evolution of the empirical share of demand covered with green offer given the optimized variables and control (empirical CFE score rho)
- Gradients of Q and $\bar{X}$ ADAM optimizers to detect when reaching optimal value
- Loss/ Cost function to make sure it is diminishing with training and get the final value which is the contract price

Here are the plots for as hyper-parameters an objective CFE score of $\rho = 0.9$, $\lambda = 100000$, a batch size of 300, $M = 70000$ epochs and the learning rate profile explicited in Figure 6. *For confidentiality purposes, I normalized the loss values (as they give out information on possible contract prices) and removed the value range on Q and $\bar{X}$ keeping only the numberless evolution with epochs.*

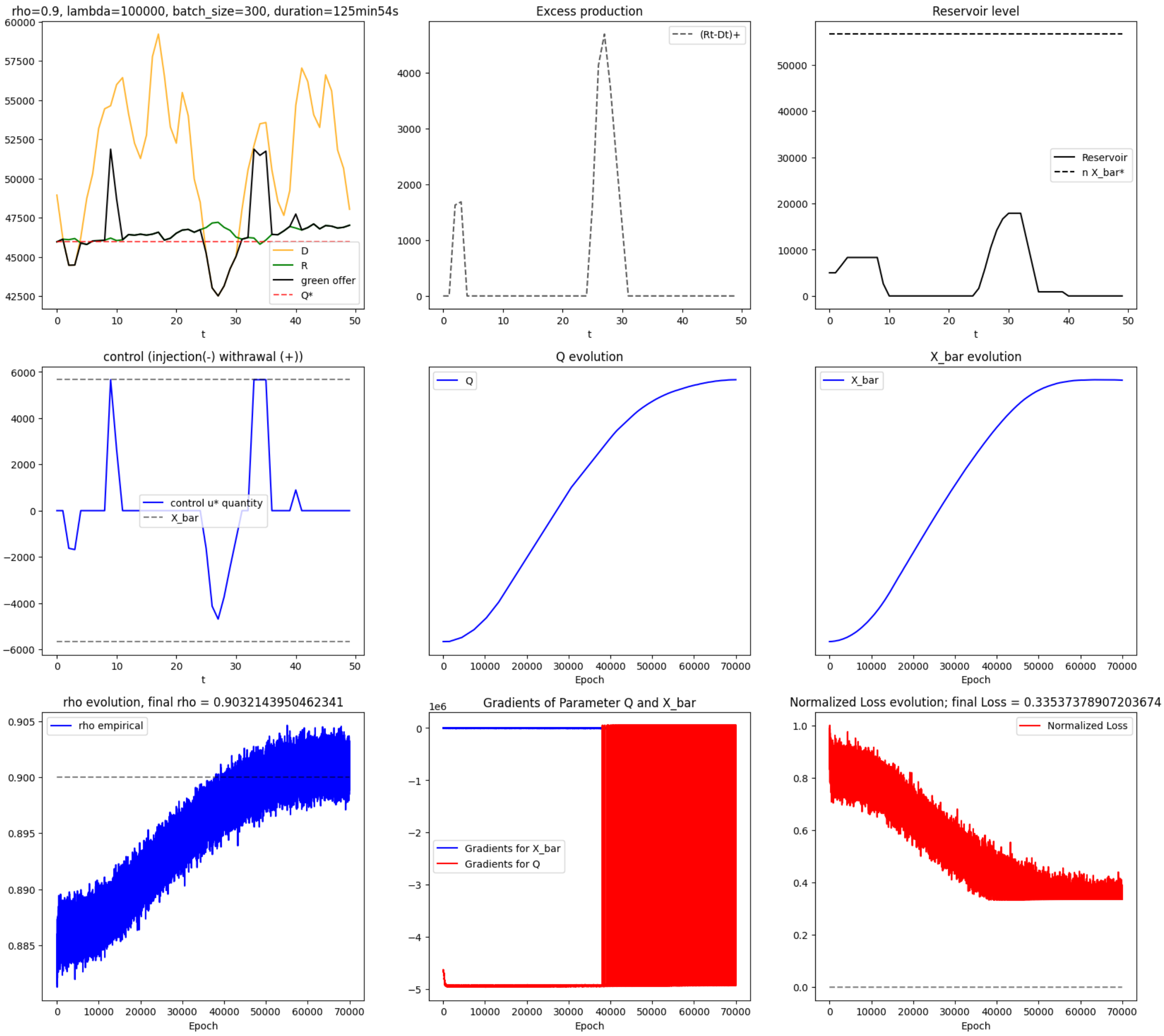

Figure 7: Analysis plots for $\rho = 0.9$ of the algorithm's results

Observations and remarks:

We can conjecture on the first plot that green offer accounts on average for around $\rho = 90\%$ of demand. At some time steps, green offer is higher than production which means there is an additional energy storage playing a role. A good thing to observe is that at all time steps, green offer never exceeds demand meaning there is no "greenwashing" as precedently warned against. Consequently, every green production at t is either used to fulfill a demand at t or stored, but never reused afterwards.

The following three plots, namely the excess production, battery level and battery storage profile confirm the use of batteries to supplement green energy production in fulfilling constraint 2, and its correlation with excess production. We can notice that battery injection and withdrawal always remain within injection and withdrawal rate $\bar{X} \times 1\text{MWh}$ (for $\bar{X}$ batteries of rate 1MWh).

On another hand, fourth and fifth graph shows us how Q and $\bar{X}$ indeed converge towards optimal values. Coherently, empirical ρ converges towards its target despite some stochasticity. Let's remind that we tried to counter stochasticity as much as computationally practical through high batch size, thus diminishing ρ's standard deviation for a fixed value of Q.

The last two graphs show us how once empirical ρ starts reaching the objective, the gradients start reacting. Indeed loss function only takes the penalized constraint if it is negative (constraint unmet), but doesn't reward meeting it. When constraint is met, loss becomes the core investment sole function $f(Q, \bar{X})$ as defined in equation 1. **The ultimate loss value is the price of the contract.**

4.2 Choosing λ

Now that we have got a converging algorithm, we need to determine which value to choose for λ. As a reminder, the λ value determines the weight or importance attributed to meeting the green constraint and appears when transforming the optimal control problem into its Lagrangian form. If the value chosen is too low, the constraint might be looked over and ρ would never reach the target. If it is too high, we risk regularization and underfitting/ suboptimal solutions.

To choose λ, we simply determine which value gives us a maximal final loss while having ρ reach the target (since it is a $\max_{\lambda \geq 0}$). In this case the contract is fulfilled and established at optimal price. The order range of f in equation 1 is around 10^8, and the penalized constraint starts around 10^3 before being multiplied by λ. We decided accordingly run the algorithm for $\lambda \in [10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8]$. If the algorithm doesn't converge (green constraint unmet), normalized loss will be set at 0. *For confidentiality purposes, loss values are normalized through softmax function:*

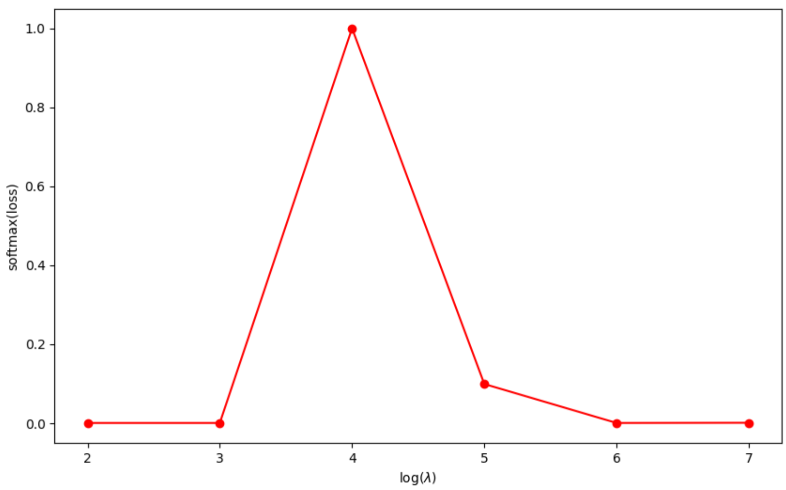

Figure 8: Loss w.r.t λ as a power of 10

The above curve clearly shows a noticeable maximum value of λ. This λ expresses the minimum penalization weight for the green constraint to be just taken into consideration, without regularization. The optimum is approximately at $\lambda = 10^4$, which will be the value we fixate on for the rest of the study.

Conclusion

In this work, we were able to set advanced foundations for the numerical resolution of a complex non-linear stochastic optimization problem with storage control. This represents a major step in the pricing of 24/7 CFE contracts as the market urgency grows rapidly because of governmental deadlines and competition with other energy firms working on the same subject.

For the rest of the internship, the new stakes are now to scale up this pricing method for it to cover the full year period. Indeed the pricing algorithm only covers 50 time steps and the contracts engage on a yearly green cover. We also will try to adjust the algorithm both in complexity and preciseness. In the most optimistic scenario, we will be able to derive a fully scaled pricing algorithm for EDF to use and commercialize its 24/7 contracts.

As machine and deep learning are currently gaining a lot of interest and use cases in quantitative finance, I am glad to be working on deep learning in a market analysis and pricing exercise. This approach is to me highly relevant as quantitative finance hedge funds nowadays rely on machine and deep learning methods to determine new and finer profit strategies. It is a way for me to stay in track of the evolution of quantitative finance with the advancement of AI and build solid ground in this field for a future career as a quantitative finance researcher.

References

[1] Andrej Karpathy. “Reproduce GPT-2 (124M)”. Youtube (2024).

[2] Aitor Lewkowycz. “How to decay your learning rate”. In: (2021).

[3] Sebastian Ruder. “An overview of gradient descent optimization algorithms”. Insight Centre for Data Analytics, NUI Galway (2017).

[4] Xavier Warin. “Reservoir optimization and Machine Learning methods”. EDF RD, FiME (2023).

[5] Xavier Warin and Olivier Bokanowski. “Neural Networks for differential games”. Université Paris-Cité, Sorbonne Univeristé, EDF RD (2024).