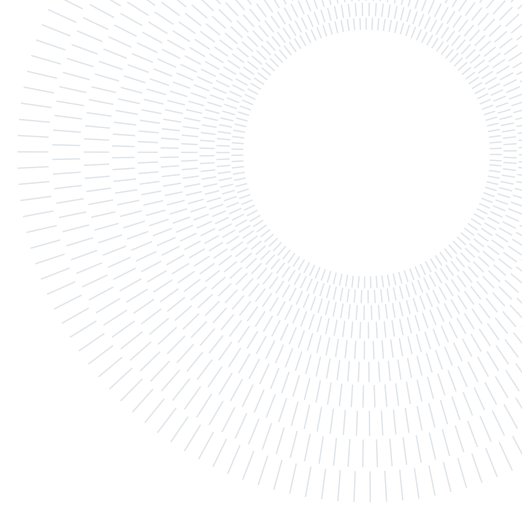# POLITECNICO
MILANO 1863

**SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE**

Multi-Objective Optimal Design of Hybrid PPA Portfolios under Uncertainty

Tesi di Laurea Magistrale in
Electrical Engineering - Ingegneria Elettrica

Lorenzo Zapparoli, 10612654

Advisor:
Prof. Filippo Bovera

Co-advisors:
Matteo Zatti

Academic year:
2022-2023

Abstract: In the recent years, governments have been progressively phasing out support schemes to renewable energy resources (RES), leading RES developers to look for alternative financing options. At the same time, due to sustainability commitments, more companies are willing to cover their demand with renewable energy and, especially after the 2022 energy crisis, they are looking for means to edge against electricity price volatility. Power purchase agreements (PPAs) can act as a synthesis of the three dynamics, and the interest towards such contracts is increasing. Bundling multi-location, multi-technology renewable and storage PPAs in hybrid portfolios has the potential of improving the financial performance while reducing the risk of PPAs investments. The difficulty in evaluating the performance and the risk of PPAs, especially when bundled in portfolios, is currently limiting their adoption. This work presents a tool to assess and design, based on buyer requirements, hybrid PPAs portfolios (including renewable and storage projects), with the objective of maximizing the financial profit while minimizing the risk from the off-taker perspective. Conditional value at risk (CVaR) is used for risk assessment, and the uncertainty on the evolution of the energy markets and the considered assets is accounted through a Monte Carlo simulation. The proposed tool is composed of three sub-models performing different tasks: long term electricity price forecasting, definition of storage operation and portfolio optimization. The three sub-models are solved in cascade to provide the final result from readily available input data. The multi-objective optimization is handled through Pareto fronts, which show the trade-off which may exist between financial return maximization and risk minimization. The optimal portfolio composition is chosen on the Pareto front depending on the risk adversity of the off-taker. The performance of the tool and the benefits of bundling multi-location, multi-technology renewable and storage PPAs are assessed on five case studies representing a menu of available contracts with different PPA prices scenarios. The results showed that the expected returns are maximized when the company invests in one or two renewable PPA projects, which are the most convenient ones; on the other hand, the inclusion of more renewable projects and storage contracts allows to reduce the risk, while reducing the expected financial performance.

Key-words: power purchase agreements, portfolio optimization, conditional value at risk, electricity price forecasting, renewable energy, ecological transition

1. Introduction

In the recent years, the drop in the levelized cost of electricity (LCOE) of renewable energy sources (RES) is making them competitive with respect to fossil fuel-based energy generation, leading governments to phase out subsidy programs [1]. Renewable projects are strongly capital-intensive, and the high volatility of market prices may make their financing risky in a purely merchant environment, therefore limiting their penetration. Consequently, renewable energy producers require alternative market mechanisms to ensure stable revenue streams and enhance financing options for RE projects. At the same, the increase in customers' awareness regarding the environmental impact of the products they buy, is leading an increasing number of companies to achieve 100% renewable energy usage as part of their sustainability initiatives. Moreover, also companies, like RE developers, look for tools for edging against the volatility of electricity prices, especially after the recent energy crisis [2, 3].

In this framework, corporate Power Purchase Agreements (PPAs), as described by [4], can be seen as the synthesis of these dynamics. Corporate PPAs are contracts between a seller and a buyer for the sale and purchase of electricity at a fixed price structure. The buyer commits to purchasing a share of the electricity generated by the seller at the PPA price, while the seller commits to procuring that same amount of electricity to the buyer. The typical duration of a PPA can range from 10 to 20 years. They introduce the benefit of transferring part of the risk related to the renewable investment from the developer to the corporate buyer, making RES projects easier to finance, while allowing the corporate buyer to get the required renewable energy certificates (RECs) and to edge against market risk. In principle this is a win-win solution, and this is demonstrated by strong increase in volume of PPAs signed in the recent years [5]. The inclusion of multiple locations and technologies in PPAs portfolios, including storage, has the potential of reducing the risk of the portfolio with respect to a single RES project, potentially making the investment in PPAs more attractive to corporate buyers. On the other side, the adoption of corporate PPAs is limited by the difficulty in evaluating the financial performance and the risk related to such contracts [6], especially when they are bundled in a portfolio and their interaction needs to be assessed.

The objective of this work is to develop a tool to determine the optimal composition of a corporate PPAs portfolio. The tool will allow the inclusion of renewable and storage PPAs, which may represent a good alternative to merchant investments for storage developers, despite not being currently widely adopted [7]. Starting from a menu of contracts, the tool defines the optimal portfolio composition, in terms of size of each contract, to minimize the risk and maximize the financial performance. The concept of conditional value at risk (CVaR) is used for the risk assessment [8].

The model is structured into three sub-models: an electricity price forecasting tool, a storage PPA modeling tool and a portfolio optimization tool. The three sub-models interact to provide the final result through a stochastic assessment starting from readily available historical data. Each of these sub-models have been investigated separately in literature [7, 8, 9]; the novelty consists in merging such models together in a single tool and modify them to fit the Italian case study, which has never been treated.

The paper is structured as follows. First, *Section 2* introduces PPAs in general, defining the considered types of renewable and storage PPAs. Then, *Section 3* describes the portfolio optimization model, defining input data and the aforementioned three sub-models. *Section 4* presents the selected case studies to assess the impact of the optimization tool on multi-technology and multi-location PPAs portfolios composition. Finally, *Section 5* draws conclusions.

2. Power Purchase Agreements

In this section the characteristics of the adopted renewable and storage PPAs are described. In general, corporate PPAs, independently from their nature, can be seen as long-term contracts between a developer (seller) and a corporate buyer for the trade of a certain amount of energy at a fixed price structure. This definition is very general and a wide variety of PPAs exist, for both storage and renewable energy trading [2]. Such level of diversification comes from a lack of standardization in the PPAs market and from the flexibility of such contracts, which allows to tune the amount of risk that is transferred from the developer to the corporate buyer. For this study just one type of renewable and storage PPA is considered, namely the virtual pay-as-produced renewable PPA and the proxy storage PPA, but the same methodology can be applied to any type of contract by changing the way in which the revenues related to the contract are modeled.

Before describing the considered contract structures, the main financial risk factors affecting PPAs performance need to be described. Considering a classic renewable PPA alone, various factors contribute to the risk (uncertainty) from the off-taker perspective [6]:

- Price risk: it derives from the uncertainty associated with electricity prices, both in terms of long-term, multi-year evolution and short-term, intra-day variability. If the future electricity price decreases below

the reference price anticipated while subscribing the PPA, the off-taker incurs smaller-than-anticipated revenues and, potentially, financial losses, in case the PPA price exceeds the future electricity price. Conversely, the off-taker realizes higher than-expected revenues if the future electricity market price is higher than the reference price.

- Volume risk: it derives by the uncertainty associated with RE generation. In particular it can reduce the off-taker revenue and undermine the possibility for the off-taker to edge sufficiently against market price risk.
- Shape risk: it derives from the interaction of electricity price and RE generation on a hourly basis. This is also known as profile or captured-price risk, with low prices occurring during moments of high generation, and vice versa. The off-taker commits to buy the generated electricity at the PPA price, but there is the risk that the hours of maximum generation coincide with the ones of minimum market price (potentially lower than the reference one), determining potentially large financial losses even if the average market price is similar or higher than the reference one [10].

Each PPA contractual structure implies a different sharing of these risks between buyer and seller. Quantifying such risk factors and their impact is crucial to determine the actual value of a PPA for the off-taker. An approach for reducing the risk of PPAs consists of combining:

- Multiple energy generation technologies, i.e. wind and solar plants, which production profiles can be considered uncorrelated [11].
- Multiple geographical locations, e.g. renewable energy projects in diverse regions and countries, for which the same technology production profiles can be considered uncorrelated.
- Renewable PPAs with storage PPAs, where storage-related revenues can be seen as complementary to solar-related ones; for example, if the prices during the day decrease, the solar PPA revenues will decrease but the proxy storage ones will increase, as the intraday price differentials will result to be increased as well.

The impact of multi-technology and multi-location PPAs on the risk factors is implicitly considered in the model through a Monte Carlo simulation, in which J scenarios of electricity prices, renewable generation and storage operation are generated to account for all the uncertainty contributions.

2.1. Virtual Pay-as-Produced Renewable PPAs

As mentioned, there are several contractual structures for renewable PPAs. In general a renewable PPA is a contract in which a RES developer and a corporate buyer agree on the trade of a determined amount of energy at a fixed price structure on a long term horizon of 10 to 20 years. The prevalent contractual structures are physical and virtual PPAs [2]. In virtual PPAs energy is not physically transferred from the generator to the off-taker; instead, the generator and the off-taker will trade energy on the wholesale electricity spot market, then the difference between the electricity spot market price and the PPA is settled separately between the off-taker and the generator, through a contract for difference (CFD) [3]. Physical PPAs, on the other hand, involve direct wiring of electricity to the off-taker's manufacturing facilities or industrial plants. This analysis focuses on virtual PPAs, but the same methodology can be applied to of physical PPAs.

It is also possible to identify different rationale to define the volumes traded in the PPA, which allow different level of risk sharing among buyer and seller. In particular two common structures are considered [8]:

- Pay-as-Produced PPA: the off-taker agrees to purchase all or a percentage of the volume produced by the RE plant, independently of its electricity demand. In this way volume risk, shape risk and price risk are fully transferred to the buyer.
- Baseload PPA: the off-taker agrees to purchase a predetermined volume of generated electricity (depending on its electricity demand). In this configuration the volume and shape risks are fully transferred to the seller through a proper contract referred as volume firming agreement (VFA), while the buyer holds the price risk.

So, in the first approach, the traded quantity will depend on the production, while in the second case it will depend on the demand.

To conclude, the considered renewable PPA contractual form is the one of pay-as-produced virtual PPA, in which the financial value of the fixed price is recognized through a CFD and the traded energy quantity corresponds to the energy produced by the renewable plant.

2.2. Proxy Storage PPAs

PPAs proved to be an effective instrument to address the major limitations of RE projects, such as high upfront investment costs and lack of guaranteed revenues. With the same limitations being faced today by energy storage technologies [12], PPAs might represent a valuable tool to foster their deployment [12]. In general, storage PPAs

are contracts signed between a corporate buyer and a storage developer in which the buyer commits to pay a certain fixed quota to the developer for a predefined period, and the developer recognizes the revenues generated by the storage on the markets, with different rationales depending on the selected contract structure. These contracts, differently from renewable PPAs, are still very rare, but are expected to increase in relevance in future years, as the levelized cost of storage (LCOS) for most technologies is expected to strongly decrease [13].

In the literature three contractual structures for storage PPAs are identified [7]:

1. Tolling agreement: this grants the buyer the right to control the storage and to operate it on multiple markets, while the seller receives energy payments to cover operational costs and capacity payments to cover fixed costs. This setup is most suited for energy traders and utilities, which have extensive expertise in energy markets, and it is applied to wind-charged storage projects in Germany [7].

2. Energy contract: here, the buyer pays a fixed price for the electricity injected by an RE plant coupled with an energy storage technology. Both the storage and the plant are operated by the seller. These contracts create an incentive mismatch, as the buyer only profits from day-ahead market revenues while the seller can also profit from offering ancillary services and performing intraday speculation. Energy contracts are currently adopted for several solar projects combined with battery storage in the US [7].

3. Proxy Storage PPA: the project developer and the energy buyer agree to sell and buy a predetermined amount of electricity (e.g. the maximum energy discharged by the energy storage in one day) at a fixed price, for a fixed time interval [7]. The cash flows for buyer and seller are:

   - From the energy buyer perspective, the revenue flows are the speculation revenues that would have been obtained in case of optimal operation of a virtual energy storage on the day-ahead market; this amount is payed by the seller to the buyer. The expense is the payment from the buyer to the seller of the fixed PPA premium.
   - From the seller perspective the revenue flows are the income due to actual use of the storage facility on the markets (day-ahead, intraday, ancillary services markets) and the fixed payment, while the expenses are the cost of the storage and the payment of the optimal day-ahead speculation revenues to the buyer.

   In other words the buyer agrees to buy each and every day a certain amount of discharge energy at the fixed PPA price; in exchange, it receives the revenues relative to the speculation of the optimally managed virtual storage in the day ahead market. The optimal management criteria have to be agreed by the parts. With this scheme the operational risk is fully transferred to the seller, and the seller can manage the storage facility however he wants. On the other side, the PPA does not restrict the operation of the storage asset in any market. Thus, it does not limit the potential market revenues for the project developer. As other revenues could potentially be achieved by the seller through the storage asset, e.g. performing intraday speculation and offering ancillary services, the total revenue for storage operators can in principle exceed the one attained via proxy storage PPA, lowering the PPA price below the LCOS of the storage and making such projects potentially more attractive for the sellers.

The selected model for this study is the one of proxy storage PPAs, being the most promising for the constitution of a corporate PPA portfolio. On the one side, the corporate buyer does not have to operate the storage, so no advanced knowledge is required, and benefits directly of the day-ahead market speculation revenues produced by the storage, thus edging on the electricity price volatility. On the other side, it allows the storage owner to operate the storage independently from the contract, allowing potentially higher revenues which could reflect in a reduction of the PPA price. Moreover, since the revenues from the buyer perspective depend only on the agreed operational scheme and on the day ahead market prices, their computation results to be simple, without the need to care about many operational constraints, as described in *Section 3*.

3. Model Description

The tool aims at determining the optimal composition of a hybrid PPAs portfolio (renewable and storage) maximizing the financial performance while minimizing the risk. The model takes as input readily available data and, adopting a stochastic approach, determines the optimal portfolio composition from the corporate buyer perspective among a menu of available contracts. The functional scheme of the model is described in Figure 1.
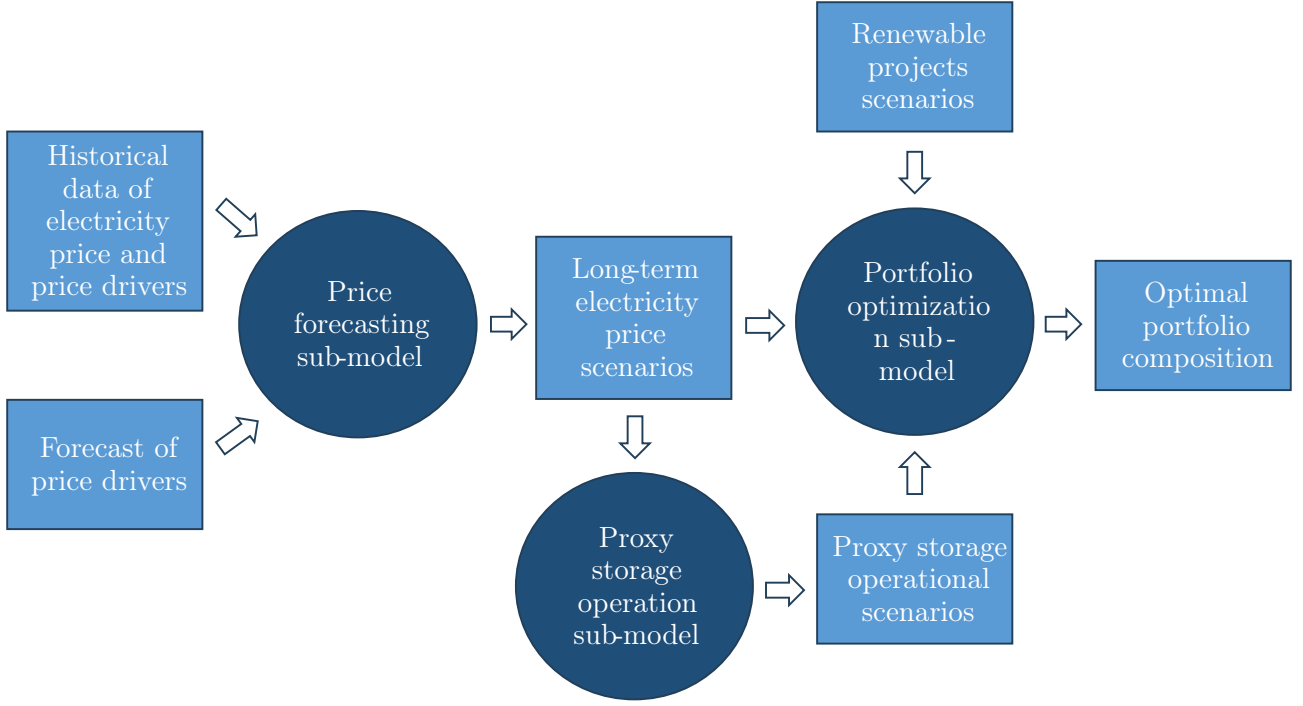
Figure 1: Portfolio optimization model functional scheme.

The tool can be seen as composed of three sub-models, each performing a different function and which are solved in cascade for the definition of the optimal portfolio composition.

1. Price forecasting sub-model: this model takes as input the historical data regarding hourly day-ahead market electricity prices, the yearly values of selected price drivers, and the forecast data regarding the selected price drivers. Then, a data-driven approach is applied using the historical data to fit a regression model correlating the price profiles to the price drivers, and applying the model the forecast of price drivers to obtain the forecast hourly electricity prices profiles over the contract duration. Given J scenarios for the future evolution of price drivers the model will provide J scenarios of the hourly future evolution of the day ahead market prices.

2. Proxy storage operation sub-model: this model takes as input the forecast electricity prices profiles and determines the operation of the storage according to the operation logic agreed between the parts. The result of this model is, for each electricity price profile, the hourly profile of the proxy storage injection over the contract duration. The injection is expressed in specific terms (MWh/MWh), as energy injected in each divided by the selected storage contract size. Given J scenarios of electricity prices the model will provide, for each proxy storage PPA, J operational scenarios.

3. PPA Portfolio optimization sub-model: given for each renewable PPA J specific hourly injection scenarios, expressed in MWh/MW (energy injected in each hour divided by the renewable contract size), and the fixed PPA price k_i, given for each proxy storage PPA the J specific injection profiles scenarios and the fixed PPA prices s_k, and given the J scenarios of the electricity prices profiles, the model determines the portfolio composition (type and size of each contract to insert in the portfolio) with two possible objectives: maximizing the expected net present value (NPV) or maximizing the conditional value at risk (CVaR). The multi objective optimization is than handled using Pareto fronts.

The input data and the algorithms adopted in each sub-model will be described in the remaining part of this section. To facilitate the understanding of each sub-model, they will be presented from the last (Portfolio optimization sub-model) to the first (Price forecasting sub-model), allowing the reader to understand first the output data each model has to provide, and only afterwards going deep into its characteristics.

3.1. Portfolio Optimization Sub-Model

The goal of the portfolio optimization sub-model is to choose the best renewable and storage PPA mix which mitigates the off-taker exposure to risk and maximizes the financial performance. The model, moving from the one presented in [8], performs the optimization with a Monte Carlo simulation in which the uncertain variables (electricity prices, renewable generation, storage operation) are introduced in J scenarios, generated according to specific rationales; then the optimal portfolio composition is obtained through a Linear Program (LP) with two objectives: maximizing the expected value of the NPV and maximize the conditional value at risk (CVaR)

of the portfolio.

3.1.1 Input Data

The input data of the multi-objective Linear Program are reported in Table 1:

Quantity	Description	Unit
$p_{t,y,j}$	Electricity price at hour t of year y in scenario j	EUR/MWh
k_i	Fixed PPA price of renewable project i	EUR/MWh
$g_{i,t,y,j}$	Specific generation of renewable project i at hour t of year y in scenario j	MWh/MW
S_i	Nominal capacity of renewable project i	MW
$q_{k,t,y,j}$	Specific injection of storage project k at hour t of year y in scenario j	MWh/MWh
s_k	Fixed proxy storage PPA price of project k	EUR/MWh
S_k^s	Nominal capacity of storage project k	MWh
r	Annual discount rate	–
Y	Contract duration	years
T	Number of hours in a year	hours
N	Number of available renewable projects	–
N^s	Number of available proxy storage projects	–
α	Confidence level adopted by the off-taker	–
D	Total off-taker demand over the contract duration	GWh
J	Number of considered scenarios	–
ϵ	Demand residual	–

Table 1: Input data.

It is possible to observe in Table 1 that for each scenario J are defined:
1. The day-ahead market price $p_{t,y,j}$ for every hour and year of the contract duration. These scenarios are generated by the price forecasting model based on the price drivers' scenarios, as it will be clarified in *Section 3.3*. The day-ahead market price is supposed the same for every contract, as the adopted price forecasting tool is not designed to capture zonal differences.
2. The specific generation of each renewable PPA project $g_{i,t,y,j}$ for every hour and year of the contract duration. Such scenarios are generated by means of random sampling of the historical time series of specific production of the considered technology in the location in which project i will be realized. Historical data are taken form online databases [14]. In other words, the future evolution scenarios of renewable generation are obtained projecting in the future the randomly-sampled historical specific production data. More advanced and accurate modeling could have been introduced to consider the expected long-term production variation applying climatic correction terms and considering the aging of the plants.
3. The specific injection of each proxy storage PPA project $q_{i,t,y,j}$ for every hour and year of the contract duration. $q_{i,t,y,j}$ assumes positive values when the storage is discharging and negative ones when it is charging. Such scenarios are generated by the proxy storage operation model, shown in *Section 3.2*, starting from the J electricity prices scenarios.

It should be noted that for each scenario j the electricity prices and storage injection profiles are fully correlated, i.e. scenario j of electricity prices is used to generate scenario j of storage injection. On the other side the renewable generation scenarios are randomly assigned to each price and storage scenario; this means that the day-ahead market price is supposed to be uncorrelated with the generation of a single specific plant. The parameters r, α, D and ϵ depend on the characteristics of the considered corporate off-taker. In particular ϵ represents the share of the demand in excess with respect to D which the portfolio can cover. For instance, if $\epsilon = 10\%$ the total renewable production contained in the portfolio can exceed D by 10 %.

3.1.2 Decision Variables

The decision variables of the optimization problem are referred as $\boldsymbol{C} \in \mathbb{R}^N$ and $\boldsymbol{C^s} \in \mathbb{R}^{N^s}$, with N being the number of available renewable projects and N^s being the number of available storage projects, as defined in Table 1. These two vectors contain all the capacities allocated to the available contracts, expressed in MW for the renewable projects and MWh for storage projects. As mentioned, the size of a renewable contract represents the contracted power, while the size of a proxy storage PPA contract represents the contracted daily discharge energy. The production (in physical value) of renewable project i in scenario j in hour t of year y can be expressed as:

$$G_{i,t,y,j} = C_i g_{i,t,y,j} \tag{1}$$

The injection (in physical value) of storage project project k in scenario j in hour t of year y can be expressed as:

$$Q_{k,t,y,j} = C_k^s q_{k,t,y,j} \tag{2}$$

Where $q_{k,t,y,j}$ is the specific proxy storage injection as defined in Table 1.

3.1.3 Problem Formulation

The portfolio optimization is performed with two objectives: maximizing the net present value and maximizing the conditional value at risk of the portfolio. First of all it is necessary to define the expression for the NPV and the CVaR in each scenario j.

The NPV for each renewable project is defined as the sum of the actualized net cash flows produced by the project over the contract duration [15], and can be expressed as:

$$NPV_{i,j} = \sum_{y=y_0}^{Y} \sum_{t=1}^{T} \frac{(p_{t,y,j} - k_i) C_i g_{i,t,y,j}}{(1+r)^{y-y_0}} \tag{3}$$

Basically, Equation 3 states that the net cash flows are defined as the difference between the market price $p_{t,y,j}$ and the fixed renewable PPA price k_i, multiplied by the production of the plant $C_i g_{i,t,y,j}$, since we are considering a pay-as-produced PPA. Then the net cash flows are actualized and cumulated to find the NPV.

Then, the NPV for each proxy storage PPA project is defined as the sum of the actualized net cash flows produced by the project over the contract duration, and can be expressed as:

$$NPV_{k,j} = \sum_{y=y_0}^{Y} \sum_{t=1}^{T} \frac{p_{t,y,j} C_k^s q_{k,t,y,j} - \dfrac{C_k^s}{24} s_k}{(1+r)^{y-y_0}} \tag{4}$$

Basically, Equation 4 states that the net cash flows for a storage contract are defined as the difference between the market price $p_{t,y,j}$ multiplied by the proxy storage injection (or withdrawal) $C_k^s q_{k,t,y,j}$ and the fixed proxy storage PPA price s_k multiplied times the amount of energy the buyer agreed to buy each and every hour within the proxy storage PPA, equal to the daily energy C_k^s divided by 24 (number of hours in a day). Then the net cash flows are actualized and cumulated to find the NPV.

Then, it is possible to sum the NPV generated by every renewable and storage project, defined in Equation 3 and Equation 4, to find the overall NPV of the portfolio in each scenario j.

$$NPV_j = \sum_{i=1}^{N} \sum_{y=y_0}^{Y} \sum_{t=1}^{T} \frac{(p_{t,y,j} - k_i) C_i g_{i,t,y,j}}{(1+r)^{y-y_0}} + \sum_{k=1}^{N^s} \sum_{y=y_0}^{Y} \sum_{t=1}^{T} \frac{p_{t,y,j} C_k^s q_{k,t,y,j} - \dfrac{C_k^s}{24} s_k}{(1+r)^{y-y_0}} \tag{5}$$

Beside the NPV, it is necessary to define the other objective used in the optimization, the conditional value at risk (CVaR), used as a measure of risk for portfolio optimization [16]. The CVaR is defined based on the confidence level that the firm is willing to accept (α), as define in Table 1, so it will be referred to as α-CVaR. Before introducing this quantity it is necessary to introduce another estimate, that is the value at risk (α-VaR), defined as the percentile of the NPV distribution with a specified confidence level (α). This parameter represents the value of NPV that has probability α of being higher than the realized NPV, or in other words the value of NPV that is exceeded with probability $1 - \alpha$ [17], as described in Figure 2. Then, it is possible to define the conditional value at risk (α-CVaR) the expected value of the NPV for values of NPV distribution lower than the α-VaR, as in Figure 2.
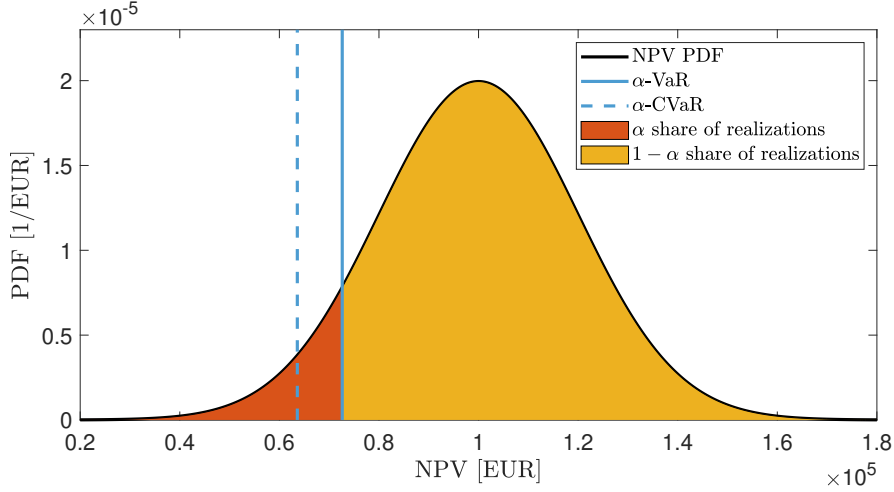
Figure 2: α-VaR and α-CVaR example.

For continuous distributions, CVaR is defined as the conditional expected NPV under the condition that NPV is lower than α-VaR [17].

$$\alpha-\text{CVaR} = \boldsymbol{E}(\mathbf{NPV}|\mathbf{NPV} < \alpha-\text{VaR}) \tag{6}$$

Let $f(\boldsymbol{x}, \mathbf{y})$ be the NPV associated to the decision variable $\boldsymbol{x} = [C, C_s]^T$ and the uncertain input data $\mathbf{y}$.

- The decision variable $\boldsymbol{x} \in \mathbb{R}^{N_x}$ can be considered ad the vector containing the share of each renewable and storage project, with N_x being the number of available projects.
- The input data $\mathbf{y} \in \mathbb{R}^{N_y}$ can be considered as the vector containing the quantities affected by uncertainty, modeled as random variables, with J being the number of available scenarios, or in other words the number of realizations.

For each $\boldsymbol{x}$, $f(\boldsymbol{x}, \mathbf{y})$ is a random variable having a distribution in $\mathbb{R}^{N_y}$ dependant on the distributions of $\mathbf{y}$. Defining the probability of a realization (scenario) of $\mathbf{y}$ as $P(\mathbf{y})$, the probability of the NPV ($f(\boldsymbol{x}, \mathbf{y})$) not exceeding a threshold ζ is:

$$\Psi(\boldsymbol{x}, \zeta) = \int_{f(\boldsymbol{x}, \boldsymbol{y}) < \zeta} P(\boldsymbol{y}) d\boldsymbol{y} \tag{7}$$

Now it is possible to define the α-VaR, that we will call $\zeta_\alpha(\boldsymbol{x})$, as the minimum value of ζ which gives probability $\Psi(\boldsymbol{x}, \zeta)$ larger or equal than α [17].

$$\zeta_\alpha(\boldsymbol{x}) = min(\zeta \in \mathbb{R} : \Psi(\boldsymbol{x}, \zeta) \geq \alpha) \tag{8}$$

As we have already said it represents the α-percentile of the NPV distribution and it is function of the decision variable $\boldsymbol{x}$.

Then, it is possible to find the conditional value at risk (α-CVaR), that we call $\phi_\alpha(\boldsymbol{x})$, applying the definition in Equation 6 [17].

$$\phi_\alpha(\boldsymbol{x}) = \frac{1}{\alpha} \int_{f(\boldsymbol{x}, \boldsymbol{y}) < \zeta_\alpha(\boldsymbol{x})} f(\boldsymbol{x}, \boldsymbol{y}) P(\boldsymbol{y}) d\boldsymbol{y} \tag{9}$$

It represents the expected value (generalized average) of the NPV distribution, under the condition that the NPV is lower than the α-VaR ($\zeta_\alpha(\boldsymbol{x})$).

As last step, since we are considering J realizations y_i of the random vector $\mathbf{y}$ in a Monte Carlo simulation, it is possible to reformulate Equation (9) in to a Linear Program (LP) by introducing the auxiliary variable $\boldsymbol{z} \in \mathbb{R}^J$, where J is the number of considered scenarios [17]. The linear formulation of the conditional value at risk is:

$$\alpha-\text{CVaR} = \phi_\alpha(\boldsymbol{x}) = \zeta - \frac{1}{\alpha} \sum_{j=1}^{J} \pi_j z_j \tag{10}$$

with

$$\begin{cases} z_j \geq \zeta - f(\boldsymbol{x}, \boldsymbol{y}_j) \\ z_j \geq 0 \\ j \in \{1, ..., J\} \\ \zeta \in \mathbb{R} \end{cases} \tag{11}$$

8

We should observe that π_j is the probability of the j-th scenario, which is supposed to be always the same, since a random sampling is performed, equal to $\pi_j = 1/J$. The new expression of the α-CVaR is simply given by the α-VaR minus the weighted average of the differences between α-VaR and NPV (loss) only for the scenarios in which NPV $< \alpha$-VaR. This is given by the first two constraints that impose that the auxiliary variable z_j must be the greater between 0 and $\zeta - f(\boldsymbol{x}, y_j)$ (the aforementioned loss); so when ϕ_α is maximized, z_j are minimized and assume value $z_j = 0$ when scenario j gives an NPV larger than α-VaR, and $z_j = \zeta - f(\boldsymbol{x}, y_j)$ when scenario j gives an NPV lower than α-VaR. The variable ζ needs to be added as decision variables.

It should be observed that maximizing the α-CVaR corresponds to moving to the right (to high values) the left tail of the NPV distribution, that is maximizing the NPV in the worst cases: this is why maximizing the α-CVaR can be seen as equivalent to minimizing the risk of the portfolio. Instead, when the expected value of the NPV is maximized, the generalized mean of the NPV distribution will tend to be maximized, but the left tail of the distribution can still assume low values, representing a higher level of risk, but average higher financial performance of the portfolio.

Once NPV and CVaR have been defined, it is possible to define the Linear Programs adopted for their maximization. First, the expected NPV maximization program is defined. The objective function to be maximized, as said, is the expected value of the NPV of the portfolio, which, starting from Equation 5, is expressed as:

$$\boldsymbol{E}\left[\mathbf{NPV}\right] = \frac{1}{J}\sum_{j=1}^{J}\left[\sum_{i=1}^{N}\sum_{y=y_0}^{Y}\sum_{t=1}^{T}\frac{(p_{t,y,j}-k_i)C_i g_{i,t,y,j}}{(1+r)^{y-y_0}} + \sum_{k=1}^{N^s}\sum_{y=y_0}^{Y}\sum_{t=1}^{T}\frac{p_{t,y,j}C_k^s q_{k,t,y,j} - \dfrac{C_k^s}{24}s_k}{(1+r)^{y-y_0}}\right] \quad (12)$$

The expected value has been converted to the average of the scenarios, since each scenarios is considered to have the same probability $\pi_j = 1/J$, as they are randomly sampled in a Monte Carlo simulation. Then, it is necessary to define the constraints, imposing at first that the demand residual from RE generation must go from zero to ϵD, as defined in Table 1. For every scenario $j \in \{1, ..., J\}$ the residual demand constraint is expressed in Equation 13.

$$0 \le \sum_{i=1}^{N}\sum_{y=y_0}^{Y}\sum_{t=1}^{T}C_i g_{i,t,y,j} - D \le \epsilon D \quad (13)$$

Then it is necessary to impose that the size allocated to each contract (renewable or storage) is bound between zero and the nominal size S_i and S_k^s, as defined in Table 1. Such constraint are represented in Equations (14), (15) and hold for every contract: $i \in \{1, ..., N\}$, $k \in \{1, ..., N^s\}$.

$$0 \le C_i \le S_i \quad (14)$$

$$0 \le C_k^s \le S_k^s \quad (15)$$

In conclusion the optimization problem can be formulated merging Equations (12), (13), (14), (15).

$$\min_{\boldsymbol{C},\boldsymbol{C}_s} \quad -\frac{1}{J}\sum_{j=1}^{J}\left[\sum_{i=1}^{N}\sum_{y=y_0}^{Y}\sum_{t=1}^{T}\frac{(p_{t,y,j}-k_i)C_i g_{i,t,y,j}}{(1+r)^{y-y_0}}\right.$$

$$\left. +\sum_{k=1}^{N^s}\sum_{y=y_0}^{Y}\sum_{t=1}^{T}\frac{p_{t,y,j}C_k^s q_{k,t,y,j} - \dfrac{C_k^s}{24}s_k}{(1+r)^{y-y_0}}\right] \quad (16)$$

$$\text{s.t.} \quad 0 \le \sum_{i=1}^{N}\sum_{y=y_0}^{Y}\sum_{t=1}^{T}C_i g_{i,t,y,j} - D \le \epsilon D \qquad j \in \{1, ..., J\}$$

$$0 \le C_i \le S_i \qquad i \in \{1, ..., N\}$$

$$0 \le C_k^s \le S_k^s \qquad k \in \{1, ..., N^s\}$$

As it is possible to see the problem is formulated as a Linear Program with $N_{tot} = N + N^s$ decision variables and $2(J + N + N^s)$ inequality constraints.

Lastly, it is necessary to define the optimization problem for α-CVaR maximization. The objective function to be maximized, as said, is α-CVaR of the portfolio, which, according to Equation 10, is expressed as:

$$\alpha-\text{CVaR} = \zeta - \frac{1}{\alpha}\sum_{j=1}^{J}\pi_j z_j \quad (17)$$

The constraints are: the ones defined in Equation (11) for the definition of the auxiliary variable, the ones defined in Equation (13) for the demand residual, the ones defined in Equations (14), (15) for the contract size

and an additional constraint for the definition of the NPV in scenario j ($f(\boldsymbol{x}, \boldsymbol{y}_j) = \text{NPV}_j$), obtained from Equation (5). The problem can be formulated as in Equation (18).

$$
\begin{aligned}
\min_{\boldsymbol{C}, \boldsymbol{C_s}, \zeta} \quad & -\zeta + \frac{1}{\alpha} \sum_{j=1}^{J} \pi_j z_j \\
\text{s.t.} \quad & z_j \geq \zeta - \text{NPV}_j && j \in \{1, ..., J\} \\
& z_j \geq 0 && j \in \{1, ..., J\} \\
& \text{NPV}_j = \sum_{i=1}^{N} \sum_{y=y_0}^{Y} \sum_{t=1}^{T} \frac{(p_{t,y,j} - k_i) C_i g_{i,t,y,j}}{(1+r)^{y-y_0}} \\
& \quad + \sum_{k=1}^{N^s} \sum_{y=y_0}^{Y} \sum_{t=1}^{T} \frac{p_{t,y,j} C_k^s q_{k,t,y,j} - \frac{C_k^s}{24} s_k}{(1+r)^{y-y_0}} && j \in \{1, ..., J\} \\
& 0 \leq \sum_{i=1}^{N} \sum_{y=y_0}^{Y} \sum_{t=1}^{T} C_i g_{i,t,y,j} - D \leq \epsilon D && j \in \{1, ..., J\} \\
& 0 \leq C_i \leq S_i && i \in \{1, ..., N\} \\
& 0 \leq C_k^s \leq S_k^s && k \in \{1, ..., N^s\}
\end{aligned}
\tag{18}
$$

As it is possible to see the problem is formulated as a Linear Program with $N_{tot} = N + N^s + 1$ decision variables, $4J + 2(N + N^s)$ inequality constraints and J equality constraints.

3.1.4 Multi-Objective Optimization

In general it can happen that the higher the potential profits, in this case the expected value of the NPV, the greater the risk of the portfolio represented by a lower value of the CVaR. In other words, when we install capacities of the RE projects to achieve the highest expected value of the NPV, the projects are simultaneously exposed to potentially lower values of NPV. This results in a reduction in the value of CVaR. To manage the multi-objective optimization, Pareto fronts [18] are used. Here two objective functions are considered: one is treated as the objective function of a single-objective optimization problem while the other is forced to take a specified value. First, a maximum NPV optimization is performed to quantify the maximum value of expected NPV, independent of risk. Then, a maximum CVaR optimization is performed to quantify the minimum portfolio's risk exposure, independent of the overall gains. The interval of CVaR is obtained with these two optimizations, adopting as lower bound the CVaR found in the NPV maximization problem and as upper bound the CVaR found in the CVaR maximization problem. This interval is divided into m intervals, corresponding to m values of CVaR. Finally, the Pareto front is built by performing m optimizations that maximize the expected value of the NPV while being subject to a constraint on the corresponding value of CVaR.

3.2. Proxy Storage Operation Sub-Model

The aim of this sub-model is to generate for each proxy storage PPA available the J scenarios of hourly specific storage injection profile $q_{k,t,y,j}$ required for the PPA portfolio optimization model as described in *Section 3.1.1*. The adopted storage PPA model is the one of proxy storage PPAs defined in *Section 2.2*, according to which the buyer agrees to buy each and every day the discharge energy C_k^s at the fixed PPA price s_k and, in exchange, it receives the revenues relative to the speculation of the optimally managed virtual storage in the day ahead market. The optimal management criterion has to be agreed by the parts, and therefore is not unique. The operational model presented in the article is inspired to the one reported in [7]. This model defines the operation of the storage as the one which maximizes the speculation revenues on the day-ahead market, subject to the main storage technical constraints. The model is described by the functional diagram represented in Figure 3.
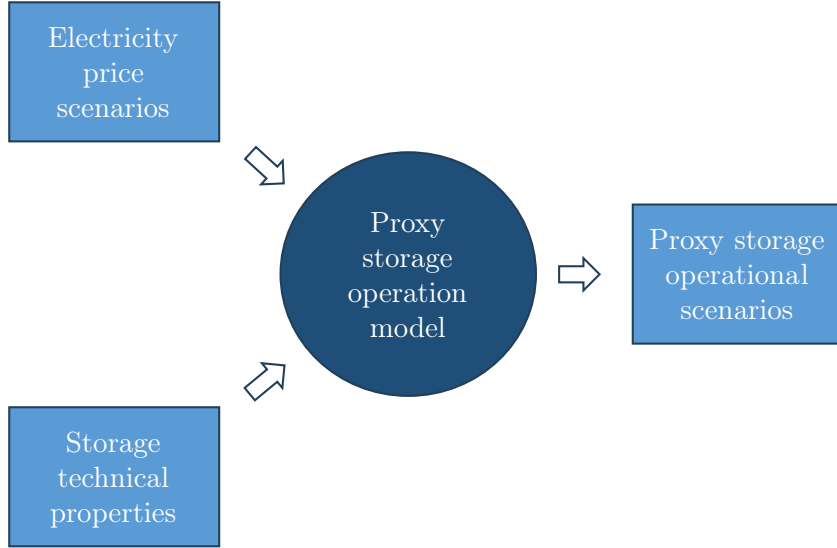
Figure 3: Proxy storage operation model functional scheme.

In particular the model takes as input the hourly electricity price profiles and the main technical characteristics of the storage, and outputs the specific injection profile according to the defined operational logic. Given J scenarios of electricity prices $p_{k,t,y,j}$ the model will provide J scenarios of storage operation. The specific injection is expressed in MWh/MWh, representing the energy injected (or withdrawn) each and every hour normalized with respect to the size of the contract, which represents the maximum injected energy allowed to the buyer (*Section 2.2*).

3.2.1 Input Data

The input data required by the model are reported in Table 2.

Quantity	Description	Unit
$p_{t,y,j}$	Electricity price at hour t of year y in scenario j	EUR/MWh
H_k	Energy-to-Power ratio of storage k	h
η_k^c	Charging efficiency of storage k	–
η_k^d	Discharging efficiency of storage k	–
n_k^c	Number of cycles per day of storage k	–
T	Number hours in a year (8760)	h
Y	Duration in years of the contracts k	y

Table 2: Input data.

Considering the input data reported in Table 2, it is possible to observe that:
1. The day-ahead market price $p_{t,y,j}$ scenarios for every hour and year of the contract duration are generated by the price forecasting model based on the price drivers' scenarios, as it will be clarified in *Section 3.3*.
2. The energy-to-power ratio H_k and the charging and discharging efficiencies η_k^c, η_k^d of every storage are defined based on the adopted technology and storage size.
3. The number of cycles per day n_c represents the number of cycles with which the storage owner intends to provide the buyer with the required energy. The concept is that, if the corporate buyer purchases 1 MWh of discharge energy in the proxy storage PPA, the seller can virtually provide it with one cycle of a 1 MWh storage or two cycles of a 0.5 MWh, and obviously the performance will be different. In other words the number of cycles per day represents how much the storage is undersized with respect to the daily energy traded in the contract.

3.2.2 Problem Formulation

The operation of each proxy storage is defined through a Linear Program. The optimization algorithm determines the optimal dispatch of the virtual storage that maximizes the revenue of the buyer (REV) over the considered time horizon (T) while fulfilling the storage constraints. The revenues coming from day-ahead speculation, i.e. the possibility of buying and selling electricity when it is most convenient, can be defined for storage k in scenario j as in Equation (19).

$$REV_{k,j} = \sum_{y=1}^{Y} \sum_{t=1}^{T} p_{t,y,j}(V_{k,t,y,j} - U_{k,t,y,j}) \tag{19}$$

$U_{k,t,y,j}$ is the charging energy of storage k in hour t of year y of scenario j and $V_{k,t,y,j}$ is the discharging charging energy of storage k in hour t of year y of scenario j. The energy for charging and discharging the virtual storage is bought and sold at the known day-ahead market price ($p_{t,y,j}$), supposed to be the same for every storage contract, as described is *Section 3.1.1*. $U_{k,t,y,j}$, $V_{k,t,y,j}$ and the state of energy $E_{k,t,y,j}$ are the decision variables of the optimization problem for all time steps; they can be grouped into vectors as: $\boldsymbol{U}_{k,j} \in \mathbb{R}^{Y \times T}$, $\boldsymbol{V}_{k,j} \in \mathbb{R}^{Y \times T}$ and $\boldsymbol{E}_{k,j} \in \mathbb{R}^{Y \times T+1}$. In total the decision variables are $3(T + Y) + 1$ for each scenario j and for each storage contract k. The additional decision variable represents the last state of energy of the storage $E_{k,T+1,Y,j}$.

The constraints are introduced to model the behavior of storage technologies. The energy stored within the storage unit in the next time step ($E_{k,t+1,y,j}$) is expressed as a linear function of the energy stored in the current time step ($E_{k,t,y,j}$), the energy charged into the storage ($U_{k,t,y,j}$), and the energy discharged from the storage ($V_{k,t,y,j}$), during the current time step, and of the charging and discharging efficiencies (η_k^c and η_k^d). Self-discharging losses are neglected, as they are negligible when operating the storage with daily cycles. Accordingly, the storage behavior is expressed through Equation (20), which applies for all time steps $t \in \{1, ..., T\}$ and $y \in \{1, ..., Y\}$.

$$E_{k,t+1,y,j} = E_{k,t,y,j} + \eta_k^c U_{k,t,y,j} - \frac{V_{k,t,y,j}}{\eta_k^d} \tag{20}$$

The stored energy, and the charging and discharging ones are non-negative quantities. Furthermore, the energy stored is constrained by the installed storage energy capacity, defined as the reference contract size divided by the number of cycles per day $E_{max,k} = C_k^{s,ref}/n_k^c$. Since the aim is to find the per-unit injection profile, $C_k^{s,ref}$ is supposed to be at the reference value $C_k^{s,ref} = 1\,\text{MWh}$. Then, the hourly charging and discharging energies are limited by the maximum charging and discharging power of the unit ($E_{max,k}/H_k$), where H_k is the equivalent duration of storage k, as defined in Table 2. These aspects result in the constraints described in Equations (21), (22), (23), which apply for all time steps $t \in \{1, ..., T\}$ and $y \in \{1, ..., Y\}$.

$$0 \leq E_{k,t,y,j} \leq E_{max,k} \Rightarrow 0 \leq E_{k,t,y,j} \leq \frac{C_k^{s,ref}}{n_k^c} \tag{21}$$

$$0 \leq U_{k,t,y,j} \leq \frac{E_{max,k}}{H_k} \Rightarrow 0 \leq U_{k,t,y,j} \leq \frac{C_k^{s,ref}}{H_k n_k^c} \tag{22}$$

$$0 \leq V_{k,t,y,j} \leq \frac{E_{max,k}}{H_k} \Rightarrow 0 \leq V_{k,t,y,j} \leq \frac{C_k^{s,ref}}{H_k n_k^c} \tag{23}$$

The next constraint aims to impose that the amount of injected energy by the virtual storage each and every day is lower or equal than the limit $C_k^{s,ref} = 1\,\text{MWh}$. This is imposed through Equation 24, for all periods $i \in \{1, ..., T/24\}$ and for all years $y \in \{1, ..., Y\}$.

$$\sum_{t=24(i-1)+1}^{24(i)} V_{k,t,y,j} \leq C_k^{s,ref} \tag{24}$$

A periodicity constraint is imposed to force the same state of charge at the beginning and at the end of given time intervals (e.g. a day); this is done to make the solution faster, as the problem can be solved independently for each day, and also not to create any energy imbalance to the storage owner. For all periods $i \in \{1, ..., T/24\}$ and for all years $y \in \{1, ..., Y\}$ this constraint is represented in Equation (25).

$$E_{k,1,1,j} = E_{k,24(i)y,j} \tag{25}$$

The energy stored at the beginning of each period ($E_{k,1,1,j}$) is not predefined but is a decision variable.

In conclusion it is possible to formulate the problem which defines the storage operation for the virtual storage k in scenario j as in Equation (26), merging the objective function defined in Equation (19) with the constraints defined in Equations (20), (21), (22), (23), (24), (25).

$$
\begin{aligned}
\min_{\boldsymbol{U_{k,j}},\boldsymbol{V_{k,j}},\boldsymbol{E_{k,j}}} \quad & -\sum_{y=1}^{Y}\sum_{t=1}^{T} p_{t,y,j}(V_{k,t,y,j} - U_{k,t,y,j}) \\
\text{s.t.} \quad & E_{k,t+1,y,j} = E_{k,t,y,j} + \eta_k^c U_{k,t,y,j} - \frac{V_{k,t,y,j}}{\eta_k^d} && t \in \{1,...,T\},\ y \in \{1,...,Y\} \\
& 0 \leq E_{k,t,y,j} \leq \frac{C_k^{s,ref}}{n_k^c} && t \in \{1,...,T\},\ y \in \{1,...,Y\} \\
& 0 \leq U_{k,t,y,j} \leq \frac{C_k^{s,ref}}{H_k n_k^c} && t \in \{1,...,T\},\ y \in \{1,...,Y\} \\
& 0 \leq V_{k,t,y,j} \leq \frac{C_k^{s,ref}}{H_k n_k^c} && t \in \{1,...,T\},\ y \in \{1,...,Y\} \\
& \sum_{t=24(i-1)+1}^{24(i)} V_{k,t,y,j} \leq C_k^{s,ref} && i \in \{1,...,T/24\},\ y \in \{1,...,Y\} \\
& E_{k,1,1,j} = E_{k,24i)y,j} && i \in \{1,...,T/24\},\ y \in \{1,...,Y\}
\end{aligned} \tag{26}
$$

This Linear Program is solved for every proxy storage PPA k and for every scenario j, obtaining the specific injection and withdrawals in hour t of year y, reported as $V_{k,t,y,j}$ and $U_{k,t,y,j}$. Lastly, the specific storage operation profile $q_{k,t,y,i}$ (considering injections as positive energy flows), is found according to Equation (27).

$$
q_{k,t,y,i} = V_{k,t,y,j} - U_{k,t,y,j} \tag{27}
$$

3.3. Price Forecasting Sub-Model

The goal of the price forecasting sub-model is to forecast the hourly day-ahead market prices profiles in different scenarios, providing for every hour t, every year y and every scenario j an estimate of the electricity price $p_{t,y,j}$ for the whole PPA contracts duration, required by the other sub-models as described in *Sections 3.1.1, 3.2.1.* Two main types of electricity prices forecasts (EPF) exist [9]:

1. Short-term EPF: in the order of hours or days, data-driven approaches are most commonly applied. Such forecasts are used typically for electricity trading and short-term operation scheduling. Such estimates are able to provide hourly resolution or more [19].
2. Long-term EPF: in the order of months or years, they are mostly obtained using econometric, market-based models [20, 21, 22, 23, 24], which capture the market dynamics by modelling the marginal cost of electricity generators. This is due to the ability of market-based models to describe, and potentially anticipate, long-term variations in the relationship between the electricity price and relevant quantities (e.g. policy-driven variations not observed in the past) [9]. Such forecasts, typically, are provided as average prices over a long time period, with low time resolution.

The adopted EPF model, inspired from the one described in [9], is data-driven and aims to provide finely-resolved long-term prediction of the wholesale electricity market price. This objective requires the combination of long-term and highly time-resolved predictions, from hourly to yearly to multi-year. The model does this starting from historical data and from market-based predictions of the most relevant quantities influencing the electricity prices profiles, namely price drivers (demand, generation, primary energy prices etc.).

This is addressed by adopting a prediction approach based on Fourier analysis, where the electricity price is described through the sum of trigonometric functions with different amplitudes and frequencies. More specifically, the hourly electricity price over a yearly horizon is decomposed into two terms:

1. Its base evolution, which is described through the dominant frequencies and is responsible for the main price behavior.
2. Its volatility, which is described by all the residual frequencies and is responsible for extreme price values.

The main idea behind this approach is that the base evolution of the electricity price can be determined by looking at the average annual values of the price drivers. The volatility of the electricity price, as well as its extreme values, are accounted for by adding a random residual profile (i.e. residual frequencies) obtained from historical hourly time series.

It is important to observe that the proposed approach does not aim at providing point forecasts of hourly market prices for electricity trading or operation optimization purposes, but rather realistic hourly price profiles for the evaluation of existing or planned assets.

The uncertainty associated with the model input data, i.e. the price drivers, is characterized performing a Monte Carlo simulation: first a probability density function is assigned to each price driver, then it is sampled to find the J scenarios of electricity prices, which are lastly translated to J scenarios of electricity prices adopting this model.

3.3.1 Input Data

The input-data set includes electricity prices, generation and demand, as well as fossil fuels and carbon prices, and aims to reflect the most relevant factors defining the electricity price. For all price drivers, two types of input data are collected:

1. Historical values, which are used to train and validate the model. Historical values of price drivers and electricity price are available from 2010 to 2021 with annual resolution; also the historical hourly values of electricity prices are available from 2010 to 2022. The sources of the data are described in Table 3.
2. Future predictions, which are used to determine future values of the electricity price. Regarding these data, several scenarios are provided by different sources. Such scenarios do not always cover all the price drivers and can span different time horizons. Here, the future prediction for the Italian energy sector are taken from the Terna scenario description 2022 [25], covering the time horizon 2022-2040.

All the input data and the relative sources are reported in Table 3.

Quantity	Unit	Historical values	Future predictions
Electricity Prices			
National Uniform Price	EUR/MWh	GME [26]	-
Total electricity sector			
Electricity demand	TWh	Terna [27]	Terna [25]
Electricity generation	TWh	Terna [27]	Terna [25]
Electricity generation mix			
Wind generation	TWh	Terna [27]	Terna [25]
Solar generation	TWh	Terna [27]	Terna [25]
Hydro generation	TWh	Terna [27]	Terna [25]
Thermal generation	TWh	Terna [27]	Terna [25]
Fossil fuel prices			
Natural gas price	EUR/MWh	BP [28]	Terna [25]
Oil price	EUR/MWh	BP [28]	Terna [25]

Table 3: Input data.

An important observation is that the future predictions of the price drivers are reported by Terna [25] in only two time shots: 2030, where two scenarios are reported ("Fit-for-55" and "Late Transition") and 2040, where three scenarios are defined ("Distributed Energy IT", "Global Ambition IT" and "Late Transition"). In 2030 the scenario "Late Transition" represents the one with lower renewable sources penetration, while the "Fit-for-55" represents the best case from the RES penetration standpoint. In 2040, the scenario "Late Transition" represents also in this case one with lower renewable sources penetration, while the "Distributed Energy IT" represents the best case. The price forecasting model, on the other side, requires as input the values of the price drivers in each and every year from 2022 to 2040, and the generation of the J price scenarios for each yearly price drivers evolution.

Considering the years 2030 and 2040, each price driver was considered to be uniformly distributed between the "best" case and the "worst" case. The assumption of an uniform distribution was made according to the principle of maximum entropy [29], not having any informative prior regarding the price drivers' distributions. Then J scenarios are created sampling the uniform distributions of each price driver in 2030 and 2040. The sampling is designed in such a way that full correlation is present between the 2030 and the 2040 realization, in other words, if sample j coincides with the "Late Transition" scenario in 2030 it will result in the "Late Transition" scenario also in 2040. In this way a set of J scenarios for each price driver in 2030 and 2040 is

created. Then an interpolation is made to obtain for every scenario j the yearly value of each price driver for all year ranging from 2022 to 2040. The results for the most relevant price drivers, as defined in *Section 3.3.2*, are depicted in the fan charts in Figure 4. It is possible to observe how the PV production and the demand are expected to increase in every scenario, while the thermal generation is expected to decrease. Moreover it is possible to observe how all the J gas price scenarios coincide: this happens because the Terna scenario [25] provides a single long term estimate of gas price, set to the same value for every scenario.

(a) PV Production.

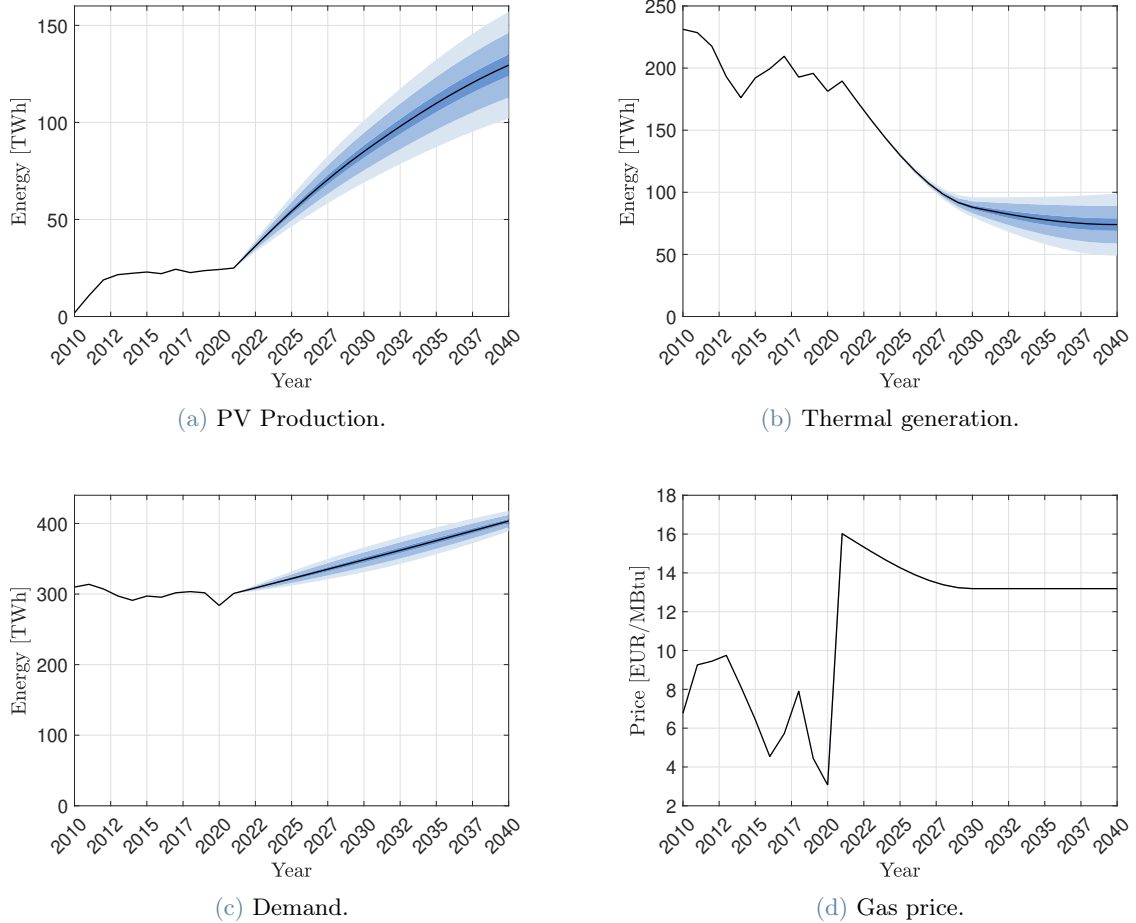(b) Thermal generation.

(c) Demand.

(d) Gas price.

Figure 4: Fan charts of the most relevant price drivers.

3.3.2 Model Description

The main idea behind the proposed approach is that a correlation exists between the base evolution of the electricity price and the average annual values of the price drivers. The base evolution of the electricity price corresponds to its most relevant fluctuations, such as annual, weekly, daily, or intra-day fluctuations.

To illustrate, consider the so-called price cannibalization effect, namely the depressive influence on the electricity price at times of high output from intermittent renewable generation. For example, high shares of solar penetration, which can be described by high values of annual solar generation, result into low values of the electricity price in summer and during the day, and into a greater seasonal and intra-day variability of the electricity price. In other words, a correlation exists between annual price drivers and the base evolution of the electricity price, which can be decomposed into a limited number of spectral components.

Following this idea, the information on the annual average values of the price drivers can be used to predict future trends of the hourly price profile. We apply Fourier analysis [30] to the hourly profiles of electricity price over an annual time horizon. In the case of the electricity price profile, the trigonometric Fourier series representation is given by Equation (28).

$$p_t = a_0 + \frac{2}{N} \sum_{n=1}^{N/2} a_n \cos\left(\frac{2n\pi t}{N}\right) + b_n \sin\left(\frac{2n\pi t}{N}\right) \tag{28}$$

The used terms are:

15

- a_0: mean value of the electricity price.
- a_n, b_n: Fourier coefficients, defining the shape of the periodic functions at frequency n.
- n: frequency (number of periods per year).
- N: length of the considered time series.
- t: time instant in which the Fourier series is evaluated.

These coefficients are determined via fast Fourier transform (FFT). Since we apply FFT on annual time series with hourly resolution, we have $N = 8760$ and $t \in \{1, ..., T\}$ with $T = 8760$ is the total number of hours in a year. The complete time series is described by a set of N coefficients, which are constant on the entire domain of the sequence, i.e. on one year. To obtain the value of the original time series, the discrete Fourier series is evaluated at time instant t. Therefore, the entire annual time series is described through a set of annual coefficients.

To obtain the base evolution of the electricity price profile, only the frequencies of the Fourier series with the largest amplitude are considered. These are denoted as main frequencies, n_{main}. Thereby, a smoothed profile capturing only the major dynamics of the electricity price is obtained. We refer to this as price approximation, which is defined in Equation (29).

$$f_t = a_0 + \frac{2}{N} \sum_{n \in n_{main}} a_n \cos\left(\frac{2n\pi t}{N}\right) + b_n \sin\left(\frac{2n\pi t}{N}\right) \tag{29}$$

f_t is the approximate (filtered) electricity price at time instant t, which consists of the main frequency components. While the price approximation captures the base evolution of the electricity price, hence the most relevant price dynamics, it is unable to describe fluctuations or times of extreme prices that go beyond the base patterns of the price profile. These are described by a residual term, R_t, which consists of all the residual frequencies not included into the price approximation. Thus, the electricity price can be expressed as in Equation (30).

$$p_t = f_t + R_t \tag{30}$$

Overall the proposed approach is based on the following steps:
1. The Fourier coefficients describing the hourly profile of historical electricity price are determined.
2. Regression models are trained by using price drivers to predict the Fourier coefficients corresponding to the price approximations.
3. Future values of the Fourier coefficients of the price approximation are predicted based on future values of the price drivers
4. Future values of the residual term are obtained by sampling historical residual term time series.
5. The future values of the Fourier coefficients are anti-transformed to obtain the base evolution in time domain, than the residual term is added to obtain the final hourly price profile forecast.

The adopted regression model takes as an input the annual values of the price drivers and produces as outputs the Fourier coefficients of the price approximation for that year. The selection of the regression model for this application is analyzed in literature [9]; in particular, the required steps are:
1. Identify the main frequencies in the historical electricity prices.
2. Select the most suitable regression model and the regressors subdividing the historical data regarding price drivers and yearly Fourier coefficients into two set, a training set and a test set, on which the performance of difference regression models are compared.

As mentioned, the first step consists in the identification of the most relevant frequency components in the Italian price profile. To find such components it is possible to apply the fast Fourier transform (FFT) to the historical price time series, reported in Figure 5a, and identify the most relevant components, including both high frequency and low frequency dynamics.

(a) Electricity price profile FFT.

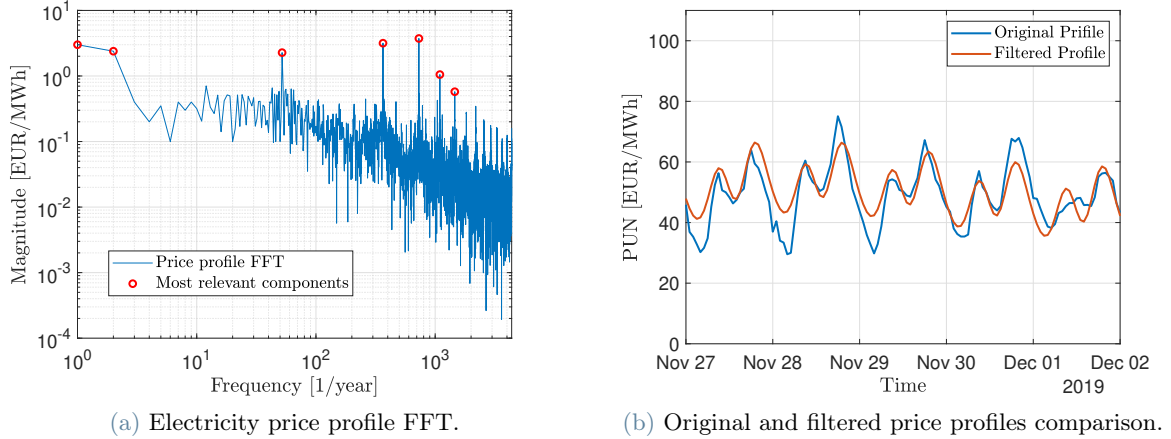(b) Original and filtered price profiles comparison.

Figure 5: Most relevant frequencies selection.

The choice was to select the eight most relevant frequency components, defined as the ones with maximum magnitude, the result is represented in Figure 5a. The selected frequencies are:

$$n_{main} = \{0, 1, 2, 52, 365, 730, 1095, 1460\}$$

Corresponding to: average price, yearly dynamics, half year dynamics, weekly dynamics, daily dynamics, half day dynamics, eight hours dynamics and six hours dynamics. The Fourier coefficients corresponding to such frequencies, according to Equation (28), are reported in Equation (31).

$$C_{main} = \{a0, a1, a2, a52, a365, a730, a1095, a1460, b1, b2, b52, b365, b730, b1095, b1460\} \tag{31}$$

Overall, the main price frequencies allow capturing the intra-day patterns given by morning and evening peaks, as well as seasonal patterns given by lower prices in summer and higher prices in winter. A comparison between the original price profile and the smoothed one, obtained just by the most relevant frequencies n_{main} is reported in Figure 5b. It is possible to observe how, as expected, the main intra-day dynamics are properly represented. The Fourier transform was applied for all years of the historical day-ahead market price time series, from 2010 to 2022, obtaining for every year the value of the Fourier coefficients relative to most relevant price frequency components, reported in Equation 31.

Once the historical data of the main Fourier coefficients are obtained, it was necessary to perform the selection of the regression model. The choice of the regression model aims to define the following aspects:

1. Which family of regression model to adopt.
2. Which historical data, among the available ones, to adopt as regressors.
3. What is the best regression model's setup for each Fourier coefficient.

An in-depth investigation of the regression model to adopt for electricity prices time series is present in literature [9]. In [9], a comparative assessment of different regression models is carried out, which includes linear regression, Gaussian process regression (GPR) and artificial neural networks (ANN). For all models, two approaches are considered:

1. Single-output: the Fourier coefficients describing the annual price approximation, i.e. a_n and b_n with $n \in C_{main}$, are predicted using separate models. The advantage of this approach stems from the potentially different relationships between different coefficients and price drivers, where all these relationship can be tuned individually.
2. Multi-output: the Fourier coefficients describing the annual price approximation are predicted using a single model. The advantage of this approach derives from the possibility of taking into account the potential interdependence of the predicted coefficients.

In the case of linear regression, various regressor configurations are investigated (linear, quadratic, and logarithmic). In the case of GPR, different kernel configurations are compared. Concerning ANN, the number of neurons and layers is varied and different activation functions are compared. All models are trained and selected using historical annual values of the price drivers and historical values of the Fourier coefficients for different European markets.

The MAPE (mean average percent error) of the price approximation function is used to assess the prediction performance of the different models. It is defined as in Equation (32).

$$MAPE = \frac{100}{H} \sum_{t=1}^{H} \frac{|\hat{f}_t - f_t|}{|f_t|} \tag{32}$$

17

where $\hat{f}_t$ s the t-th prediction of function f and f_t its observed value.

The result, according to literature [9], is that the best regression model is Gaussian process regression (GPR), in the single-output configuration. For this reason, the decision was to adopt GPR as a regression model also in the Italian case study, supposing the correlation between price drivers and price profile features to be similar in European countries. Under this assumption, fifteen GPR models need to be defined, one for each Fourier coefficient.

Once the regression model to adopt is identified, it is necessary to define for each GPR model the kernel function and the regressors to adopt. The selection of such features was performed dividing the historical data (price drivers and Fourier coefficients) into two set: the training set, ranging form 2010 to 2019 and the test set consisting of year 2020. The regression models are trained on the training set and their performance is compared on the test set, adopting the MAPE as a key performance indicator (KPI), as defined in Equation (32). It was chosen to discard the data relative to 2021 and 2022 because in these years, due to the energy crisis, the long term (yearly or seasonal) dynamics of the price drivers were determined more by the intra-year variation of the price drivers (e.g. gas price) than the yearly average value of the aforementioned drivers. This leads the models relative to the lower frequency dynamics $(a1, b1, a2, b2)$, to be wrongly trained. For this reason it was chosen to discard years 2021 and 2022 from the evaluation and training of the models.

The selection of the regressors (price drivers) and the kernel functions was performed in an iterative fashion. First, fixing the kernel functions and finding the price drivers which, used as regressors, provided the lowest MAPE; then, fixed the regressors, finding the best kernel functions, and so on till the convergence. The kernel functions that were investigated are: linear and squared exponential. Instead the investigated regressors are the price drivers reported in Table 3. The result of the model tuning are reported in Table 4.

Fourier coefficient	Kernel function	Regressors
$a0$	Linear	Gas Price, PV production, Thermal production, Demand
$a1$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$a2$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$a52$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$a365$	Linear	Gas Price, PV production, Thermal production, Demand
$a730$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$a1095$	Linear	Gas Price, PV production, Thermal production, Demand
$a1460$	Linear	Gas Price, PV production, Thermal production, Demand
$b1$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$b2$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$b52$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$b365$	Linear	Gas Price, PV production, Thermal production, Demand
$b730$	Squared exponential	Gas Price, PV production, Thermal production, Demand
$b1095$	Linear	Gas Price, PV production, Thermal production, Demand
$b1460$	Linear	Gas Price, PV production, Thermal production, Demand

Table 4: Kernel functions and regressor choice.

The configuration of kernels and regressors defined in Table 4 is the one which provides the lowest value of MAPE. In particular it results that MAPE = 20 % for 2020 price profile, with models trained on 2010 to 2019 profiles. The actual 2020 price profile and the forecast filtered profiles can be visually compared in Figure 6.

(a) 2020 Forecast and actual price profiles.

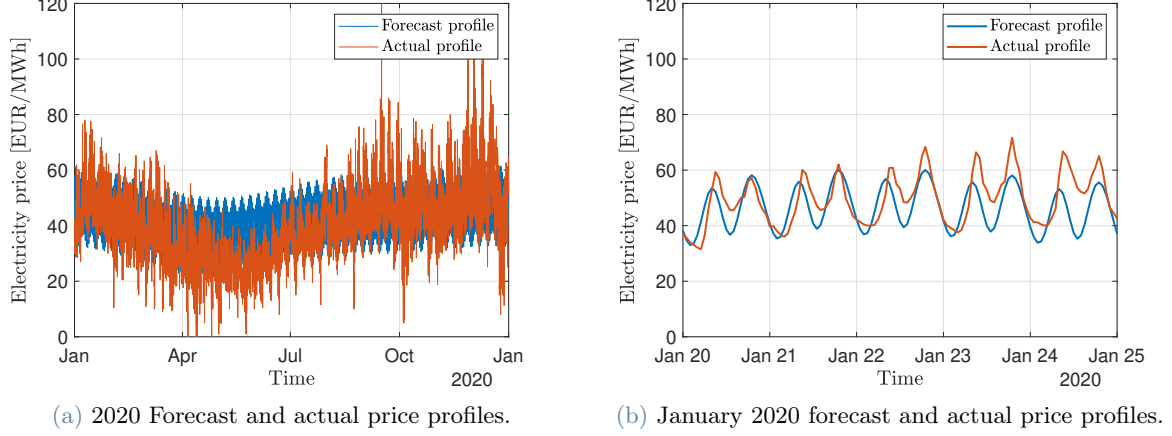(b) January 2020 forecast and actual price profiles.

Figure 6: Forecast and actual price profiles comparison.

In particular, in Figure 6a the overall yearly profiles are compared, and it is possible to observe how the regression model is capable of describing the main features. The same observation can be extended to Figure 6b, in which the daily price profiles are compared; in particular, it is possible to observe how the model is capable also to capture the main intra-day variations of the electricity price.

Once the regression models have been selected and validated, they can be applied to the scenarios of price drivers described in *Section 3.3.1* to obtain the scenarios of the most relevant Fourier coefficients. Then the scenarios of the Fourier coefficients are transformed into filtered electricity prices scenarios f_t, through Fourier anti transform. Lastly the historical residuals R_t are randomly sampled and added the forecast filtered electricity prices, as in Equation (30), to obtain the forecast electricity prices p_t for every scenario j. The forecast is performed for every scenario j and for every year y to obtain the prices scenarios $p_{t,y,j}$, required as input data from the other models, as described in *Section 3.1.1, 3.2.1*.

3.3.3 Results

In this section the forecast electricity prices scenarios are described. As mentioned in *Section 3.3.1*, the regression model is trained on the historical data from 2010 to 2020 and then is applied to the price drivers scenarios to find the hourly electricity prices scenarios from 2021 to 2040 over a 20 years horizon. A total of $J = 100$ scenarios were generated. The results are represented in Figure 7.

(a) Forecast electricity prices box plot.
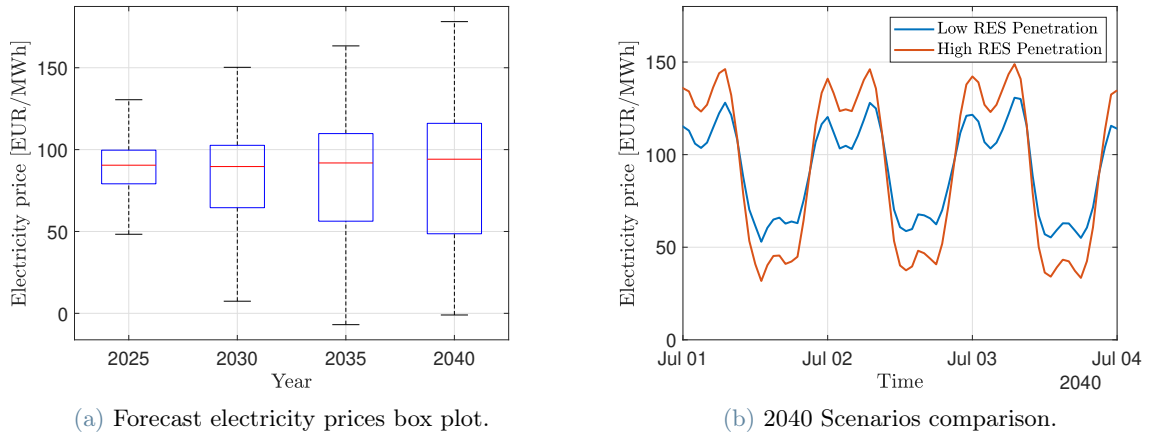

(b) 2040 Scenarios comparison.

Figure 7: Forecast electricity prices scenarios.

Considering the box plot in Figure 7a, the red line in each box represents the average electricity price in the considered year, the blue lines represent the 25th and 75th percentile and the black bars the extreme values assumed by the data.

A first observation from Figure 7a is that the average electricity price tends to stay the same from 2025 to 2040; this can be explained by the fact that the average electricity price can be seen as correlated to the gas price

and, in the price drivers definition described in *Section 3.3.1*, the gas price is supposed to be the same in 2030 and 2040, as from Terna hypotheses [25]. Then, always considering Figure 7a, it is possible to observe how the electricity price distribution tends to be more and more dispersed. This happens because the intra-day price differentials between day and night tend to increase with the passing of time. The reason in that, according to the price drivers scenarios, shown in Figure 4, the PV production will tend to increase with time to reach the environmental objectives, and this causes the day-time prices to decrease and the night-time prices to increase due to the cannibalization effect. The cannibalization effects arises when there are many renewable plants which production is fully correlated: when they produce all together they cause a depression of the price due to the drastic increase in the offer. With PV this effect is particularly evident, since all PV plant produce in day-time and take energy from the same primary source.

Figure 7b represents the electricity day-ahead market price forecasting in July 2040 in two scenarios, defined in *Section 3.3.1*: "Late Transition", which here for the sake of clarity is called "Low RES Penetration", and "Distributed Energy IT", which here for the sake of clarity is called "High RES Penetration". These two scenarios are the extremes of the uniform distribution defined in *Section 3.3.1*. It is possible to observe how the "High RES Penetration" scenario the prices during the day tend to be lower and the prices during the night tend to be higher than the "Low RES Penetration" scenario. The increase of the price differentials at the increase of the PV penetration is again caused by the cannibalization effect.

Considering all the observations on Figure 7a, 7b, it is possible to conclude that the price forecasting model is capable to capture the main effect of the price drivers on the electricity price profiles, like the cannibalization effect and the correlation between the average electricity price and the gas price.

4. Simulations and Results

In this section the PPA portfolio optimization tool will be tested on a set of case studies to assess it's validity and appreciate the advantages of bundling portfolios of PPAs containing multi-location, multi-technology and storage contracts. Moreover it will be possible to appreciate the trade-off between NPV maximization and CVaR maximization through Pareto fronts.

The menu of contracts among which the portfolio optimization tool will be called to choose consists of a set of eight renewable contracts and three storage contracts, modeling the main technologies and locations on the Italian territory. Five different scenarios will be simulated varying the prices of each class of PPAs to perform a sensitivity analysis of the optimal portfolio composition with respect to the PPA prices.

The scenarios are described in *Section 4.1*, and the portfolio optimization results for each are shown and commented in *Section 4.2, 4.3, 4.4, 4.5, 4.6*.

4.1. Assumptions and Scenarios Description

In this section the main assumptions and scenarios adopted for the different simulations will be described. In particular the aim is to define all the input parameters required by the different sub-models, as described in Table 1, 2, 3. In particular, two input data classes are defined: the ones that are the same in all the simulated scenarios and the ones which change from one scenario to the other, namely the PPA prices. First, the selected contract menu will be described, then the characteristics of the selected corporate buyer and lastly the selected PPA prices scenarios adopted for the sensitivity analysis.

4.1.1 PPA Contract Menu

All the simulations share the same menu of renewable and storage PPAs, among which the tool can select the contracts for the portfolio composition. The choice was to adopt $N = 8$ renewable PPAs and $N^s = 3$ storage PPAs, representative of different generation and storage technologies and of different locations in the Italian territory.

In the selection of the renewable projects the choice was to select four utility-scale PV projects and four wind projects, two on-shore and two off-shore. This was done to introduce in the menu all the main renewable technologies which are expected to be developed in Italy. Regarding the localization of each project, the choice was made looking at the cumulated power of connection requests presented to the Italian transmission system operator (TSO) [31], placing the projects in those places which show the maximum amount of connection-requested power for each selected technology. This placement logic aims to locate the considered contracts in the places where most of the new generation is expected to be developed. One exception was made: the logic would have selected all the four PV projects in the south, but it was chosen to introduce also a PV plant in the north of Italy, to stress the multi-location nature of the menu. Such location was selected as the one with the

maximum connection-requested power in the north of Italy. The selected locations are reported in Table 5 and in Figure 8.

Contract name	Technology	Location (Region)
PV Contract 1	Solar	Catania (Sicily)
PV Contract 2	Solar	Foggia (Apulia)
PV Contract 3	Solar	Sassari (Sardinia)
PV Contract 4	Solar	Voghera (Lombardy)
Wind Contract 1	Wind onshore	Foggia (Apulia)
Wind Contract 2	Wind onshore	Sassari (Sardinia)
Wind Contract 3	Wind offshore	Lecce (Apulia)
Wind Contract 4	Wind offshore	Trapani (Sicily)

Table 5: Renewable projects menu.

(a) PV location map.
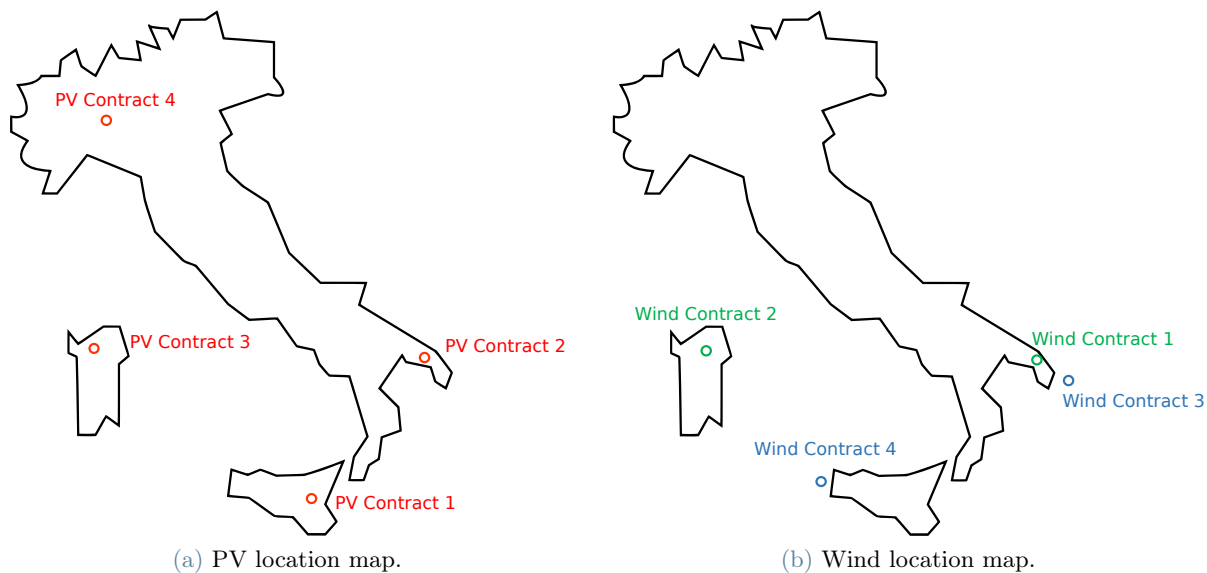
(b) Wind location map.

Figure 8: RES projects location maps.

For each selected location, the specific production historical data were taken from [14] using the MERRA-2 database. In particular, for the solar projects single-axis tracking was supposed with an azimuth of 180 deg and a balance of system efficiency of $\eta_{bos} = 90\,\%$. For wind projects a standard Vestas V90 2000 turbine with hub height of 80 m was adopted. All the plants were supposed of unity size of 1 MW to obtain the specific historical production profile. Then, for each contract, the historical profiles were randomly sampled and projected in the future, as described in *Section 3.1.1*, to obtain $J = 100$ scenarios of specific hourly renewable production scenarios $g_{i,t,y,j}$ required as input to the portfolio optimization sub-model, as in Table 1. The box plots of the equivalent hours of each renewable project is reported are Figure 9, in which the red line in each box represents the average equivalent hours (h_{eq}) in the considered year, the blue lines represent the 25th and 75th percentile and the black bars the extreme values assumed by the data.
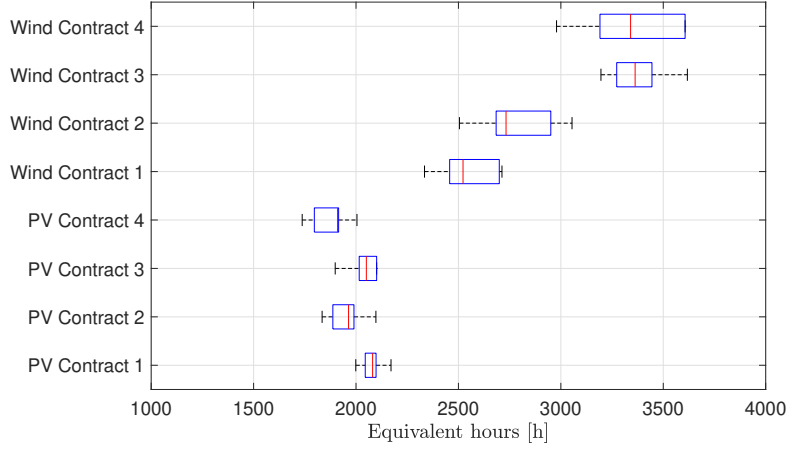
Figure 9: Equivalent hours of the renewable projects menu.

Considering the proxy storage PPA contracts, the choice was to select three contracts representative of the two main currently available storage technologies: electrochemical storage (lithium battery based) and pumped hydro. The parameters to define for each proxy storage contract k are reported in Table 2: the energy to power ratio H_k, the charging efficiency η_k^c, the discharging efficiency η_k^d and the number of cycles per day n_k^c. The selected proxy storage contracts are defined in Table 6. The data were found in literature [22].

Contract name	Technology	H_k	n_k^c	η_k^c	η_k^d
Storage Contract 1	Electrochemical	4 h	2	95 %	95 %
Storage Contract 2	Electrochemical	8 h	1	95 %	95 %
Storage Contract 3	Pumped hydro	8 h	1	80 %	80 %

Table 6: Storage projects menu.

It was chosen to select two electrochemical storage contracts to model the fact that a storage developer can provide to the corporate buyer the agreed daily discharge energy in two ways: with a storage of the same size used for one cycle per day or with a storage of half size used for two cycles per day. Clearly the two approaches will result in a different storage operation profile, and therefore in a different value of the contract. The technical parameters are set as input to the proxy storage operation sub-model, described in *Section 3.2*, to obtain the J scenarios of storage operation $q_{k,t,y,j}$ required as input by the portfolio optimization sub-model, as in Table 1. The performance of the three storage contracts is represented through the box plots in Figure 10.

(a) Storage contracts daily specific revenue.

(b) Evolution in time of Storage Contract 2 daily specific revenue.
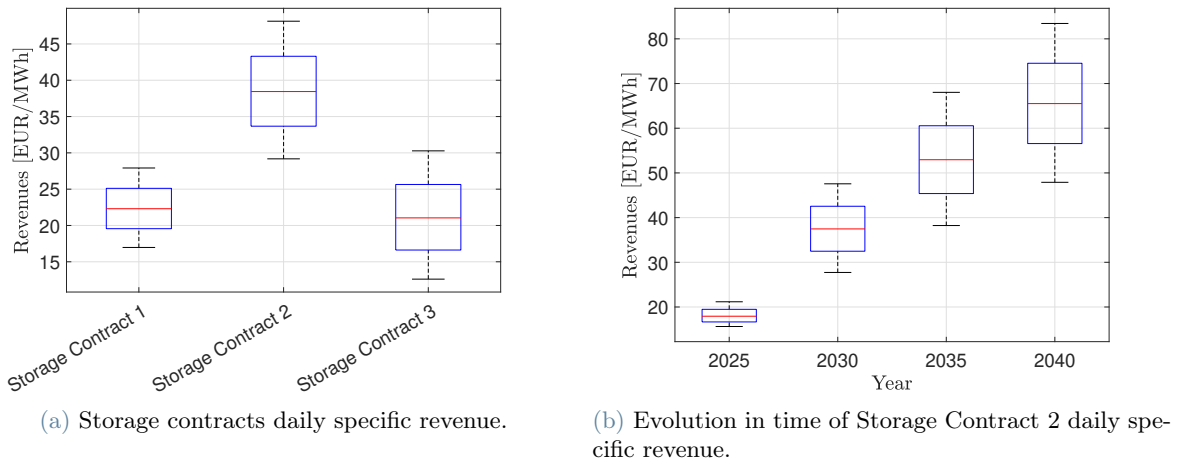
Figure 10: Proxy storage contracts properties.

The box plot in Figure 10a, expresses the distributions of daily speculation revenue produced by each storage contract, expressed in specific term (EUR/MWh) with respect to the contract size. It is possible to observe how Storage Contract 2 is capable of producing the highest daily revenues. This is explained by the fact that Storage Contract 1, despite being the same technology and providing the injected energy every day, is half the capacity, so the captured price differentials are reduced. On the other side Storage Contract 3 has the same capacity of Storage Contract 2, but has sensibly lower efficiency, due to the technological differences.

The box plot in Figure 10b represents the evolution in time of the specific daily revenues distribution of Storage Contract 2. It is observed that the revenues increase drastically over time, following the increase in daily price differentials denoted in Figure 7a. Comparing the PPA price with the evolution of the daily specific revenues, it is possible to understand at which point in time the storage starts to generate revenues: this happens when the daily revenues exceed the PPA price.

The nominal (maximum) size of each contract is set to a reference value of $S_i = 10\,\text{MW}$ for each renewable contract and $S_k^s = 10\,\text{MWh}$ for each proxy storage contract. The contract duration for every storage and renewable PPA is supposed to be the same and equal to $T = 20\,\text{years}$, form 2021 to 2040. This value is reasonable for renewable and pumped-hydro storage PPAs, as their useful life is longer than the contractual duration. On the other side, the useful life of an electrochemical storage is typically lower than 20 years [13]: the hypothesis is that, keeping the same contract, the storage developer revamps the battery when it is needed, allowing to sign contracts for a time longer than the useful life of the specific plant. This option comes with no additional risk for the corporate buyer, since its revenues are not correlated to the actual operation of the storage, as described in *Section 2.2*.

4.1.2 Corporate Buyer Attributes

The attributes of the corporate buyer, required as input by the portfolio optimization sub-model, according to Table 1 are:

1. D: demand of the corporate buyer over the contract duration of $Y = 20\,\text{years}$.
2. r: annual discount rate adopted by the firm.
3. α: confidence level adopted by the off taker. This parameter is required for the computation of the CVaR, as described in *Section 3.3*. It represents the adversity to risk of the corporate buyer, the higher α, the higher the accepted risk, and vice versa.
4. ϵ: demand residual. It represents the maximum share of the demand D, by which the portfolio overall renewable production can exceed the demand D over the contract duration.

The values adopted for such parameters are reported in Table 7.

Quantity	Value
D	$100\,\text{GWh}$
r	$4\,\%$
α	$10\,\%$
ϵ	$10\,\%$

Table 7: Corporate buyer attributes.

The choice of the parameters in Table 7 depends on the considered corporate buyer. In particular, typical values have been selected for medium-sized industry, as typical of the Italian territory.

4.1.3 PPA Prices Scenarios

In *Scetion 3.3.3, 4.1.1, 4.1.2* all the input data required in Table 1 have been defined, excluding the prices k_i of each renewable PPA and s_k of each storage PPA. The definition of the contracts' prices will be described in this section. The price selection is performed generating five scenarios:

1. The first four scenarios adopt standard PPA prices found in literature [32], and consist in the four combinations of "high" and "low" renewable PPA prices and "high" and "low" storage PPA prices scenarios. These four scenarios aim to investigate the effect of different renewable and storage PPA prices on the portfolio composition.
2. The fifth scenarios defines the PPA prices based on the LCOE calculation for each plant. This scenario allows to model the effect of the location (north-south) and the technology (onshore-offshore) on the PPA prices, and therefore on the portfolio composition.

At first, the first four scenarios will be described. In particular, the renewable PPA prices are described in the LevelTen Energy 2023 Report [32]. This report defines the average PPA prices signed in different European countries, divided by year (from 2020 to 2023) and by technology (PV or Wind). It should be observed that PPA prices for wind are not defined for Italy in the report, because no wind PPA was signed in the considered period; for this reason, France was considered as benchmark for wind projects. Moreover, the report does not make any differentiation between onshore and offshore wind, so the price has been supposed to be the same for both technologies. Regarding PV, average data for Italy are available, the PPA price is supposed to be the same for all the contracts defined in Table 5. It should be observed that current PPA prices are inflated by the recent energy prices, so 2023 PPA prices will be referred to as the high renewable PPA price scenario. On the other side, 2020 prices represent the pre-crisis prices, and will be referred as the low renewable PPA price scenario. The data reported in [32] are summarized in Table 8, 9.

Contract	Technology	PPA Price k_i	Year
PV Contract 1, 2, 3, 4	Solar	73 EUR/MWh	2023
Wind Contract 1, 2, 3, 4	Wind (onshore and offshore)	100 EUR/MWh	2023

Table 8: High renewable PPA price scenario.

Contract	Technology	PPA Price k_i	Year
PV Contract 1, 2, 3, 4	Solar	44 EUR/MWh	2020
Wind Contract 1, 2, 3, 4	Wind (onshore and offshore)	77 EUR/MWh	2020

Table 9: Low renewable PPA price scenario.

Considering, instead, proxy storage PPAs, no historical data are available, so the PPA price was supposed to be equal to the levelized cost of storage (LCOS) of the three considered projects, as suggested in [7]. In literature [13], the values of LCOS for the three considered technologies are reported as 2017 values and 2030 projections. Regarding 2030 projections three cases for each technology are considered: "best case", corresponding to the lowest LCOS; "base case", corresponding to the average LCOS; and "worst case", corresponding to the highest LCOS. For this analysis, "best case" values are always adopted, in order to model a scenario in which storage prices drastically drop, eventually thanks to subsidies [7]. Storage PPA prices (LCOS) in 2030 are expected to be substantially lower than the ones in 2017; for this reason 2030 prices will be referred to as the low storage PPA price scenario and 2017 prices will be referred as high storage PPA price scenario. The data regarding proxy storage prices (LCOS) are summarized in Table 10, 11.

Contract	Technology	PPA Price s_k	Year
Storage Contract 1	Electrochemical	140 EUR/MWh	2017
Storage Contract 2	Electrochemical	190 EUR/MWh	2017
Storage Contract 3	Pumped hydro	80 EUR/MWh	2017

Table 10: High storage PPA price scenario.

Contract	Technology	PPA Price s_k	Year
Storage Contract 1	Electrochemical	27 EUR/MWh	2030
Storage Contract 2	Electrochemical	35 EUR/MWh	2030
Storage Contract 3	Pumped hydro	80 EUR/MWh	2030

Table 11: Low storage PPA price scenario.

The two scenarios of renewable PPA prices and the two scenarios of storage PPA prices are combined, as described in Table 12, to produce the four scenarios of prices for the whole menu of available contracts.

	High renewable PPA price	Low renewable PPA price
High storage PPA price	Scenario 1	Scenario 2
Low storage PPA price	Scenario 3	Scenario 4

Table 12: Simulated scenarios.

The portfolios resulting from the PPA prices scenarios in Table 12 are described in the following sections.
Lastly, scenario five needs to be described. In the previous scenarios the renewable PPA prices are supposed to be the same for all the projects of the same technology, in this scenario, instead, a price differentiation is introduced to model the effect of location and technology on the renewable PPA prices. The PPA prices were calculated computing the levelized cost of electricity (LCOE) of the various projects and finding the PPA price adding a risk premium of 10 EUR/MWh for each. The LCOE was computed as described in [33], starting from the capital and operational expenditures indicated for each technology in the IRENA report [34] for the European market and the average yearly equivalent utilization hours of each project, represented by the red bars in Figure 9. The results are reported in Table 13.

Contract	Technology	PPA Price k_i
PV Contract 1	Solar	52 EUR/MWh
PV Contract 2	Solar	51 EUR/MWh
PV Contract 3	Solar	51 EUR/MWh
PV Contract 4	Solar	53 EUR/MWh
Wind Contract 1	Wind onshore	73 EUR/MWh
Wind Contract 2	Wind onshore	70 EUR/MWh
Wind Contract 3	Wind offshore	91 EUR/MWh
Wind Contract 4	Wind offshore	92 EUR/MWh

Table 13: Scenario 5 renewable PPA prices.

Lastly, in Scenario 5 the low storage PPA price scenario, detailed in Table 11, is adopted for proxy storage prices.

4.2. Scenario 1

As described in Table 12, scenario 1 is characterized by high renewable PPA prices and high storage PPA prices, and is representative of the 2023 market. The optimization of the portfolio among the eleven available contracts led to the selection of 2.73 MW of PV Contract 2 in case of expected NPV maximization, and led to the selection of 2.61 MW of PV Contract 2 in case of CVaR maximization. Since the two single-objective optimizations provide different result, the multi-objective optimization is handled through a Pareto front, as described in *Section 3.1.4*. The resulting Pareto front and the histograms of the two single-objective optimizations are depicted in Figure 11.
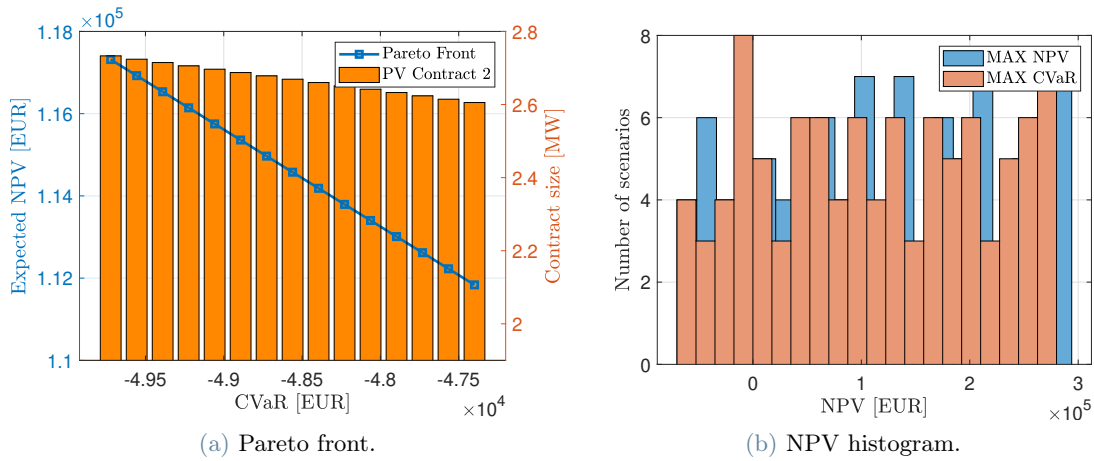
(a) Pareto front.

(b) NPV histogram.

Figure 11: Scenario 1 portfolio optimization results.

In Figure 11b it is possible to observe the difference between the distribution of the $J = 100$ scenarios of NPV resulting from the two single-objective optimized portfolios. In particular the histogram relative to NPV maximization shows a higher value of the expected NPV, as the distribution is slightly shifted to the right, but also shows more samples in the left tail indicating a lower value of CVaR. On the other side the CVaR maximization histogram shows a lower expected NPV with the advantage of having less samples in the left tail, indicating higher CVaR and lower risk.

Considering the Pareto front in Figure 11a, it is clearly possible to see the trade-off between CVaR and expected NPV. Moving on the front, at the increase of the CVaR the expected NPV tends to decrease and vice-versa. The corporate buyer has to select a spot on the front depending on its risk adversion. The increase in CVaR, in this case, is given by a slight reduction in the selected contract power for PV Contract 2.

In Scenario 1, the optimization tool composed a portfolio just with one renewable contract in all cases(PV Contract 2): no storage or portfolio differentiation was introduced. It is worth observing that PV Contract 2 is not the solar contract with highest equivalent utilization, as it can be observed from Figure 9, but the model selects it anyways as the best contract also for NPV maximization. This may seem counter-intuitive, but the leading concept is that in a PPA the buyer pays the produced energy, not the installed power; for this reason, if the PPA price is the same for all PV contracts, the portfolio optimization tool will select just the contract with the profile which captures the highest prices, in this case being PV Contract 2. Clearly, the required power to cover the corporate demand will be higher than the case in which PV Contract 1 were selected, but this make no difference as long as the PPA energy price is the same. The effect of the equivalent hours of the plant typically reflects on the PPA price, plants with higher equivalent hours show lower prices and vice-versa, this effect is modelled in Scenario 5. Storage contracts are not selected because financially inefficient, as it is possible to understand comparing the proxy storage PPA prices of this scenario with the storage-related revenues in Figure 10.

4.3.   Scenario 2

As described in Table 12, scenario 2 is characterized by low renewable PPA prices and high storage PPA prices. The optimization of the portfolio among the eleven available contracts led to the selection of 1.41 MW of PV Contract 1 and 1.26 MW of PV Contract 2 in both cases of expected NPV and CVaR maximization. The resulting Pareto front and histograms of the two single-objective optimizations are depicted in Figure 12.

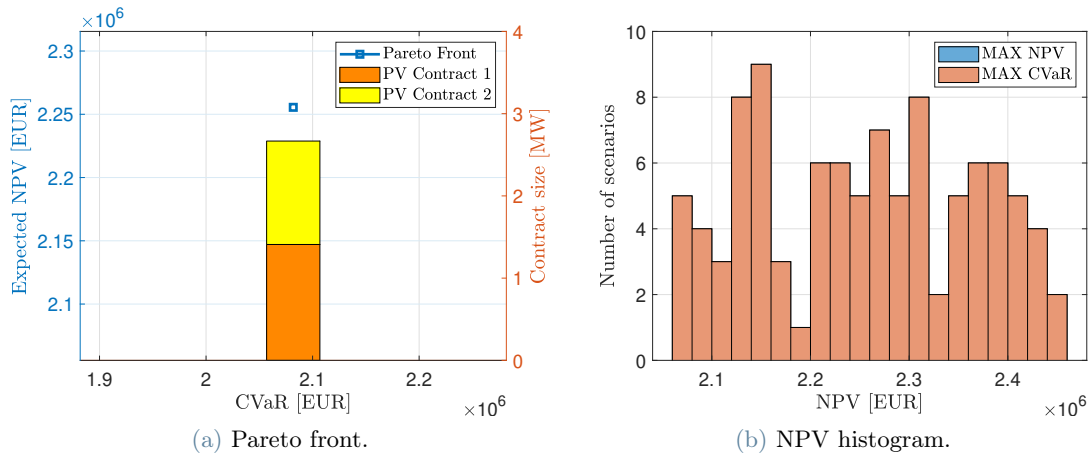(a) Pareto front.

(b) NPV histogram.

Figure 12: Scenario 2 portfolio optimization results.

As it is possible to observe in Figure 12a, since the two single-objective optimizations provide the same result, the Pareto front degenerates into a point, and the two NPV histograms coincide, as in Figure 12b. This means that the obtained portfolio configuration, composed of the mix of two contracts, is capable at the same time to maximize the expected value of the NPV and minimize the risk.

As in Scenario 1, storage contracts are not selected because financially inefficient.

4.4.   Scenario 3

As described in Table 12, scenario 3 is characterized by high renewable PPA prices and low storage PPA prices. The optimization of the portfolio among the eleven available contracts led to the selection of 2.73 MW of PV Contract 2 in the case of expected NPV maximization and 1.61 MW of PV Contract 1, 1.05 MW of PV Contract 2 and 4.65 MWh of Storage Contract 2 in case of CVaR maximization. The resulting Pareto front and histograms of the two single-objective optimizations are depicted in Figure 13.
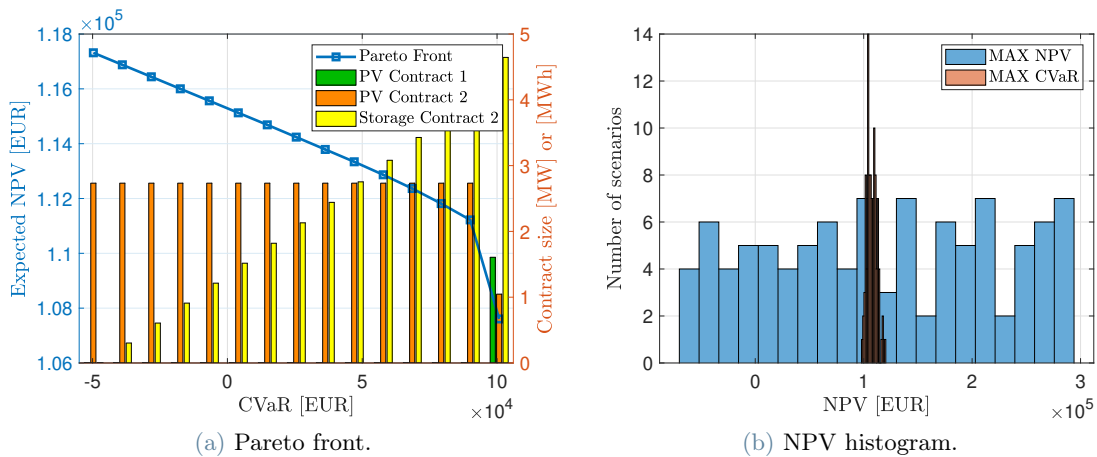

(a) Pareto front.

(b) NPV histogram.

Figure 13: Scenario 3 portfolio optimization results.

Considering Figure 13b, it is possible to observe how the dispersion of the histogram relative to the CVaR maximization is lower than the one relative to the expected NPV maximization, resulting in a much higher CVaR and a lower risk of the portfolio. This result is obtained mainly through the inclusion of storage in the PV portfolio, which allows to edge against captured price risk. In the price scenarios in which the intra-day price differentials are stronger, due to cannibalization effect, the revenues generated by the PV contracts is lower, while the one related to the storage contract is higher; vice-versa, when price differentials are low, the revenues of PV contracts increase and the ones of storage contracts decrease. These two opposite effects allow to reduce the dependence of the NPV on the price differentials, this way reducing the related risk.

As it is possible to observe in Figure 13b, PV Contract 1 is included in the portfolio only in the last point of the Pareto front, determining a strong decrease in NPV for just a slight increase in CVaR. For this reason the global optimal solution could be the last-but-one point in the front, in which only Storage Contract 2 and PV

27

Contract 2 are selected.

In this scenario, differently form the previous one, the storage is selected, as its prices are decreased and are now compatible with the generated revenues. Among the three contracts, Storage Contract 2 was selected because it is the one which produces the highest specific daily revenues, as in Figure 10a, and this compensates the lower PPA price of Storage Contract 1.

An interesting analysis consists in the evaluation of the maximum storage PPA price (threshold price) for which each storage contract is included in the maximum expected NPV portfolio. This threshold price represents the maximum storage PPA price for which each storage is financially convenient on its own; in other words it produces alone a positive NPV. Then, if no storage is included in the maximum NPV portfolio, while some are included in the maximum CVaR one, it is possible to compare each considered proxy storage price with its threshold price and understand the value given to storage's risk-reduction effect in the CVaR maximization portfolio. The threshold prices computed by the model are reported in Table 14.

Contract	Treshold PPA Price
Storage Contract 1	20 EUR/MWh
Storage Contract 2	34 EUR/MWh
Storage Contract 3	17 EUR/MWh

Table 14: Threshold storage PPA prices.

As expected Storage Contract 2 has the maximum threshold price, as it is the one able to produce the maximum daily revenues. If we compare such values with the low storage PPA price for Storage Contract 2, it is possible to see that the price is 35 EUR/MWh, so 1 EUR/MWh higher than the threshold price: this explains why Storage Contract 2 in not included in the maximum expected NPV portfolio, and indicates that, in the portfolios in which the storage is included, a premium of 1 EUR/MWh is payed to achieve storage's risk reduction effect.

4.5. Scenario 4

As described in Table 12, scenario 4 is characterized by low renewable PPA prices and low storage PPA prices. The optimization of the portfolio among the eleven available contracts led to the selection of 1.41 MW of PV Contract 1 and 1.26 MW of PV Contract 2 in the case of expected NPV maximization and 1.54 MW of PV Contract 1, 1.12 MW of PV Contract 2 and 4.62 MWh of Storage Contract 2 in case of CVaR maximization. The resulting Pareto front and histograms of the two single-objective optimizations are depicted in Figure 14.

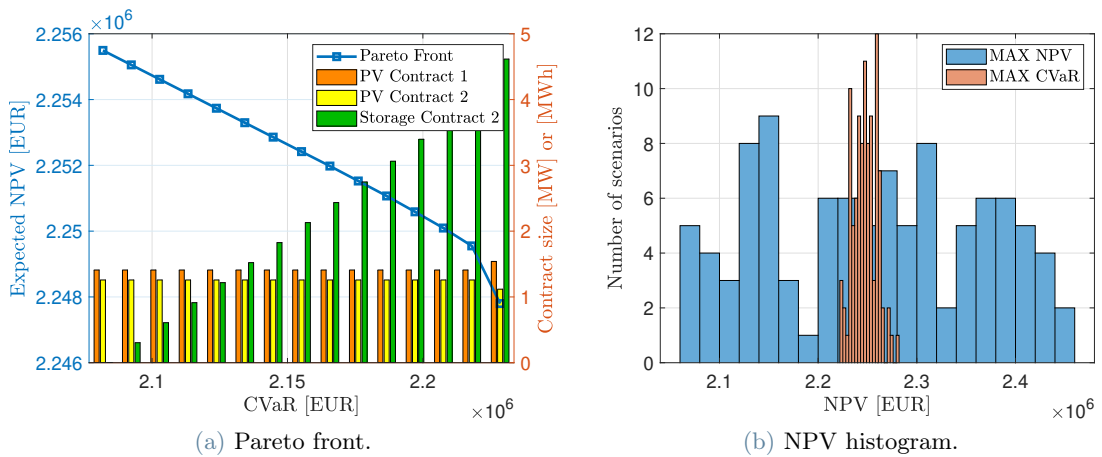(a) Pareto front.

(b) NPV histogram.

Figure 14: Scenario 4 portfolio optimization results.

In this case the selection of renewable contracts is the same as in scenario 2, with which scenario 4 shares the renewable PPA prices. Then, observing Figure 14a, 14b, the low storage prices determine the inclusion of Storage Contract 2 in the portfolio at the increase of the CVaR, with the consequent reduction in the NPV distribution dispersion, as observed for Scenario 3.

4.6. Scenario 5

As described in Table 12, scenario 5 is characterized by low storage PPA prices and renewable PPA prices obtained through the LCOE of the singe projects. The optimization of the portfolio among the eleven available contracts led to the selection of 1.89 MW of Wind Contract 2 in the case of expected NPV maximization and 1.89 MW of Wind Contract 2 and 0.31 MWh of Storage Contract 2 in case of CVaR maximization. The resulting Pareto front and histograms of the two single-objective optimizations are depicted in Figure 15.

(a) Pareto front.

(b) NPV histogram.
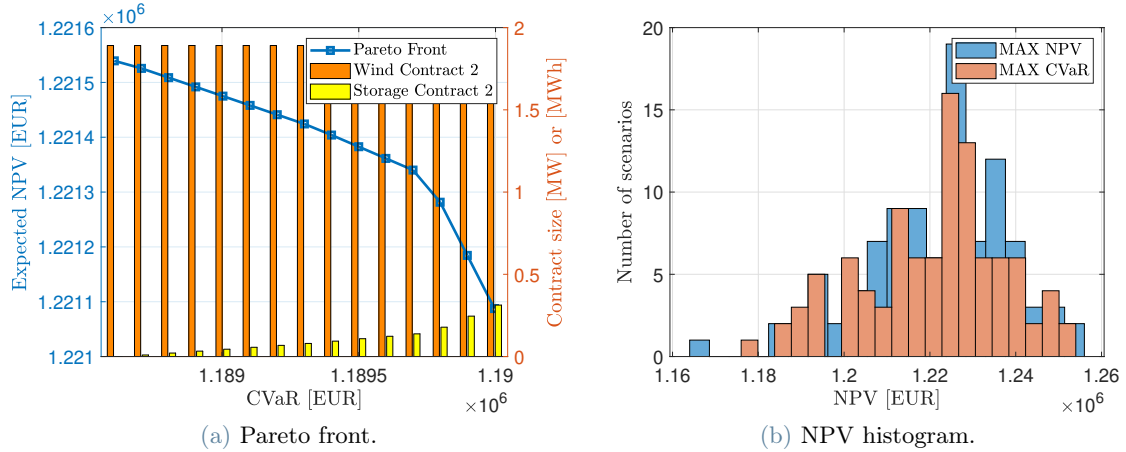
Figure 15: Scenario 5 portfolio optimization results.

In this scenario the renewable project which results to maximize the expected NPV and the CVaR is Wind Project 2; this because in this PPA prices scenario the price differential between PV projects and onshore wind projects is reduced. Moreover, as it can be observed in Figure 15a, an increase in CVaR is achieved including Storage Contract 2 in the portfolio. Differently from the previous scenarios, the inclusion of storage does not have a big impact on the NPV distribution (Figure 15b). This can be explained considering that the storage contract here is coupled with a wind project, so the cannibalization effect is less relevant and a daily-operated storage has little means to hedge against it.

5. Conclusion

This work presents a tool for the optimization and assessment of PPA portfolios under uncertainty, including multiple generation and storage technologies. The uncertainty is treated through a Monte Carlo simulation and the optimal portfolio composition is defined through a multi-objective optimization, aiming to maximize the expected revenues and minimizing the risk. The two objectives are investigated through two different objective functions: expected NPV for profit maximization and CVaR for risk minimization. The multi objective optimization is carried out building Pareto fronts in order to capture the trade-off between the different objectives. The corporate buyer has to decide the portfolio composition as the one corresponding to the point on the front best suiting its risk adversion. The proposed tool is tested on five different case studies, modeling different PPA prices scenarios in the available contract menu.

We find that, in most of the cases, the portfolios resulting from revenues maximization and risk minimization are different, meaning that a trade-off between expected revenues and risk is typically present. In other words, if the corporate buyer is willing to reduce the risk of the portfolio, it has to accept a reduction in the expected revenues, which can be seen as a premium payed for risk mitigation, originated by the inclusion of inefficient, but risk-reducing, contracts in the portfolio. The portfolios which maximize the expected NPV are typically composed of just renewable contracts and the level of differentiation is scarce, typically one contract, maximum two, in each portfolio. The portfolios which maximize the CVaR, instead, can show the inclusion of storage and a higher level of differentiation. This means that storage contracts on their own are not convenient, but their coupling with renewable contracts (especially solar ones) introduces a strong risk reduction effect which may make their inclusion a viable option, depending on the risk adversion on the buyer. Risk mitigation is also achieved, as in scenario 3, including multi-location contracts in the portfolio. In the considered case studies a multi-technology portfolio (PV and Wind) is never observed, but this does not prove in general that technology diversification does not bring any advantage in terms of risk mitigation, this just means that the considered PPA prices scenarios and the selected contract menu do not create an opportunity for technological diversification. The study also shows that storage, at the current LCOS, is never included, because it is too economically

inefficient to justify its volatility mitigation effect. In the future, LCOS of the main storage technologies is expected to decrease, and the possibility guaranteed by proxy storage PPAs to the storage developer to freely operate the storage in any market and keep the revenues, has the potential to reduce the storage PPA prices below the LCOS. In such scenarios the inclusion of storage in PPA portfolios may become a viable option, and storage PPAs may become a valuable solution for financing and fostering the spread of storage systems.

To summarize, the developed tool allows the corporate buyer to design and evaluate PPA portfolios under uncertainty. It allows to capture the value and the risk associated to each contract and is capable of assessing the effect of bundling more contract in portfolios. These capabilities are exploited to design the portfolio which best suits the requirements of the corporate buyer. The modularity of the tool makes it scalable and prone to improvements, e.g. it is possible to replace the data-driven price forecasting sub-model with a potentially more accurate econometric model, or it would be possible to select different kind of PPA contractual structures modifying just the way in which their revenues are computed. In conclusion, the purpose of this work is to contribute to the energy transition proposing a tool that company can use to design a tailored PPA portfolio, which would allow the decarbonization of their demand while producing a positive economical return.

References

[1] IEA. Renewable energy market update 2021. Technical report, IEA, Paris, 2021. Available at `https://www.iea.org/reports/renewable-energy-market-update-2021`. (Accessed 10 May 2023).

[2] RE-Source. Introduction to corporate sourcing of renewable electricity in europe. Technical report, Re-Source Platform, 2020. Available at `https://resourceplatform.eu/`. (Accessed 20 May 2023).

[3] Yashar Ghiassi-Farrokhfal, Wolfgang Ketter, and John Collins. Making green power purchase agreements more predictable and reliable for companies. *Decision Support Systems*, 144:113514, 2021.

[4] James Kobus, Ali Nasrallah, and Jim Guidera. The role of corporate renewable power purchase agreements in supporting us wind and solar deployment. *Center on Global Energy Policy, Columbia University SIPA, March*, 24, 2021.

[5] Pexapark. European ppa market outlook 2023. Technical report, Pexapark, 2023. Available at `https://pexapark.com/european-ppa-market/?creative=617204887582&keyword=european%20ppa%20market%20outlook%202023&matchtype=e&network=g&device=c&utm_campaign=MLT_Market-Outlook_OS-PRA_TLA_MKTO_WARM_SRCH_CON_MC_ALL_CPC_ONG_ONG&gclid=Cj0KCQjw84anBhCtARIsAISI-xeqqKn0qmqSPJkgT6W6tbJ4HO3fGWOVxqy3rhdx_TTiWRuyim22Sp8aApy_EALw_wcB`. (Accessed 23 May 2023).

[6] RE-Source. Risk mitigation for corporate renewable ppas. Technical report, Re-Source Platform, 2020. Available at `https://resourceplatform.eu/`. (Accessed 20 May 2023).

[7] Paolo Gabrielli, Philipp Hilsheimer, and Giovanni Sansavini. Storage power purchase agreements to enable the deployment of energy storage in europe. *iScience*, 25(8), 2022.

[8] Paolo Gabrielli, Reyhaneh Aboutalebi, and Giovanni Sansavini. Mitigating financial risk of corporate power purchase agreements via portfolio optimization. *Energy Economics*, 109:105980, 2022.

[9] Paolo Gabrielli, Moritz Wüthrich, Steffen Blume, and Giovanni Sansavini. Data-driven modeling for long-term electricity price forecasting. *Energy*, 244:123107, 2022.

[10] Javier López Prol, Karl W Steininger, and David Zilberman. The cannibalization effect of wind and solar in the california wholesale electricity market. *Energy Economics*, 85:104552, 2020.

[11] Stefan Pfenninger and Iain Staffell. Long-term patterns of european pv output using 30 years of validated hourly reanalysis and satellite data. *Energy*, 114:1251–1265, 2016.

[12] Lindsay Miller and Rupp Carriveau. A review of energy storage financing—learning from and partnering with the renewable energy industry. *Journal of Energy storage*, 19:311–319, 2018.

[13] Martin Beuse, Bjarne Steffen, and Tobias S Schmidt. Projecting the competition between energy-storage technologies in the electricity sector. *Joule*, 4(10):2162–2184, 2020.

[14] Imperial College London and ETH Zürich. *Renewables Ninja*, 2023. Availabe at `https://www.renewables.ninja/` (Accessed 23 May 2023).

[15] Chenghui Tang and Fan Zhang. Classification, principle and pricing manner of renewable power purchase agreement. In *IOP Conference Series: Earth and Environmental Science*, volume 295, page 052054. IOP Publishing, 2019.

[16] R Tyrrell Rockafellar, Stanislav Uryasev, et al. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

[17] Pavlo Krokhmal, Jonas Palmquist, and Stanislav Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of risk*, 4:43–68, 2002.

[18] Patrick Ngatchou, Anahita Zarei, and A El-Sharkawi. Pareto multi objective optimization. In *Proceedings of the 13th international conference on, intelligent systems application to power systems*, pages 84–91. IEEE, 2005.

[19] Florian Ziel and Rick Steinert. Probabilistic mid-and long-term electricity price forecasting. *Renewable and Sustainable Energy Reviews*, 94:251–266, 2018.

[20] Zaid Mohamed and Pat Bodger. Forecasting electricity consumption in new zealand using economic and demographic variables. *Energy*, 30(10):1833–1843, 2005.

[21] Antonio Bello, Derek Bunn, Javier Reneses, and Antonio Muñoz. Parametric density recalibration of a fundamental market model to forecast electricity prices. *Energies*, 9(11):959, 2016.

[22] Antonio Bello, Javier Reneses, Antonio Muñoz, and Andrés Delgadillo. Probabilistic forecasting of hourly electricity prices in the medium-term using spatial interpolation techniques. *International Journal of Forecasting*, 32(3):966–980, 2016.

[23] Florentina Paraschiv, David Erni, and Ralf Pietsch. The impact of renewable energies on eex day-ahead electricity prices. *Energy Policy*, 73:196–210, 2014.

[24] Andreas Coester, Marjan W Hofkes, and Elissaios Papyrakis. Economics of renewable energy expansion and security of supply: A dynamic simulation of the german electricity market. *Applied Energy*, 231:1268–1284, 2018.

[25] Terna. Documento di descrizione degli scenari 2022. Technical report, Terna, 2022. Available at `https://download.terna.it/terna/Documento_Descrizione_Scenari_2022_8da74044f6ee28d.pdf`. (Accessed 28 May 2023).

[26] Gestore dei Mercati Energetici. *GME - Download - dati MGP - prezzi*, 2023. Available at `https://www.mercatoelettrico.org/it/Download/DownloadDati.aspx?val=MGP_Prezzi`. (Accessed 12 June 2023).

[27] Terna. *Pubblicazioni Statistiche*, 2023. Available at `https://www.terna.it/it/sistema-elettrico/statistiche/pubblicazioni-statistiche`. (Accessed 10 June 2023).

[28] BP. Statistical review of world energy. Technical report, BP, 2022. Available at `https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html`. (Accessed 10 June 2023).

[29] Peter Harremoës and Flemming Topsøe. Maximum entropy fundamentals. *Entropy*, 3(3):191–226, 2001.

[30] Peter Bloomfield. *Fourier analysis of time series: an introduction*. John Wiley & Sons, 2004.

[31] Terna. *Econnextion platform*, 2023. Available at `https://www.terna.it/it/sistema-elettrico/rete/econnextion.aspx?val=MGP_Prezzi`. (Accessed 20 June 2023).

[32] LevelTen Energy. Ppa price index executive summary. Technical report, LevelTen Energy, 2023. Available at `https://go.leveltenenergy.com/l/816793/2023-07-19/37bxrx/816793/1689802704o9nNagit/2023Q2_EU_Executive_Summary_PPAPriceIndex.pdf`. (Accessed 20 July 2023).

[33] Kadra Branker, MJM Pathak, and Joshua M Pearce. A review of solar photovoltaic levelized cost of electricity. *Renewable and sustainable energy reviews*, 15(9):4470–4482, 2011.

[34] IRENA. Renewable power generation costs in 2021. Technical report, IRENA, 2021. Available at `https://www.irena.org/publications/2022/Jul/Renewable-Power-Generation-Costs-in-2021`. (Accessed 21 July 2023).