POLITECNICO DI TORINO

Master's Degree in Mathematical Engineering

Master's Degree Thesis

Optimal Sizing of Renewable Energy Communities Under Uncertainty

Supervisors
Prof. Lorenzo Bottaccioli
Dott. Daniele Salvatore Schiera

Candidate
Aurelio Zizzo

Academic Year 2023-2024

Abstract

In recent years, the global landscape of energy production and consumption has witnessed a shift towards more sustainable and community-driven models. Renewable energy communities (RECs) have emerged as dynamic and participatory entities, where individuals, businesses, and local organizations unite to collectively generate, manage, and share energy resources. This transformation turns them from passive consumers into active participants, or prosumers, contributing to the transition towards clean energy and fostering decarbonization. Realizing a REC demands substantial investments in renewable energy sources (RES), and the economic feasibility of such projects is pivotal for their successful establishment and long-term sustainability. Recognizing this, institutions are actively pushing for their diffusion adopting incentivizing policies. In the Italian context, RECs are promoted through the valorization of the shared energy, a performance indicator for diffused self-consumption configurations, strongly related to the synchronism between generation and consumption. This synchronism, in a pre-feasibility study, is often difficult to evaluate due to the strong variability in renewable energy sources production and the users' consumption patterns, leading to an unreliable assessment of the economic feasibility of the project. Within this paradigm, this thesis aims to develop a stochastic optimization model to determine the optimal design of the energy production plant from renewable sources, specifically photovoltaic (PV) technology, to maximize the Net Present Value (NPV) of the project, and to evaluate the possibility of the installation of an aggregate storage system within the REC. The problem is formulated as a two-stage stochastic MILP with recourse, where first-stage decision variables account for the sizing of PV and storage systems, and second-stage variables optimize energy dispatch operations. Uncertainty is introduced into the model following a probabilistic approach based on Monte Carlo simulations: different scenarios are generated, starting from historical data, in order to capture the variability of the uncertain parameters and to consistently represent their possible combinations. The uncertainty modelling mainly involves the photovoltaic generation, the energy load, and energy prices. The hourly PV production is modelled with a Gaussian mixture distribution and daily PV profiles are reconstructed by sampling sequentially from the fitted random variables. A similar procedure is applied to generate load profiles and energy price scenarios. The impact of each stochastic parameter on the output is analyzed and stopping criteria for Monte Carlo simulations are evaluated to ensure a tradeoff between accuracy and computational times. Economic and energetic Key Performance Indicators (KPIs), including payback time, self-sufficiency, and self-consumption, are then computed to assess the goodness and feasibility of the solution. Finally, the

results of the stochastic problem are tested and compared against the deterministic solution, obtained by solving the problem considering the average values of the uncertain parameters. The solution to the stochastic problem can lead, depending on the configuration of the CER, to an increase of up to 15% in NPV, resulting in a drastic reduction of the payback time of the investment.

Table of Contents

List of Tables

List of Figures

Introduction

In the 21st century, the world is facing an unprecedented challenge that transcends geopolitical boundaries and threatens our planet's ecosystems: climate change. The escalating concentration of greenhouse gases in the Earth's atmosphere, primarily carbon dioxide (CO2), is the basis of this challenge, stemming from the combustion of fossil fuels and contributing to the greenhouse effect and subsequent global warming. Figure 1 illustrates the exponential increase in global CO2 emissions from energy combustion and industrial processes in recent decades, with a growth of 0.9% or 321 Mt in 2022, reaching a new all-time high of 36.8 Gt. [1].

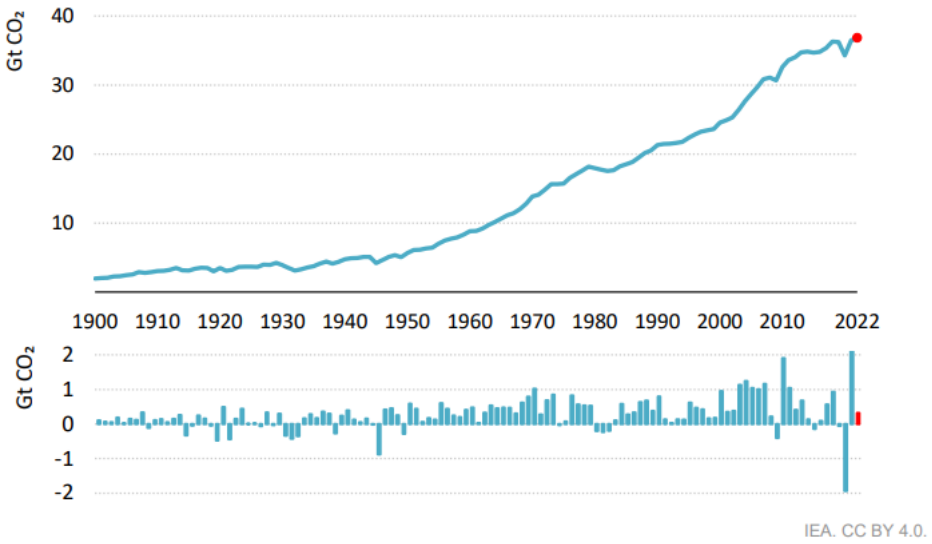

Figure 1: Global CO2 emissions from energy combustion and industrial processes and their annual change, 1900-2022.

Decarbonisation is crucial in mitigating the negative effects of CO2 emissions. This involves reducing carbon intensity across sectors, transitioning from fossil fuels to cleaner energy alternatives, and adopting innovative technologies to capture and

store carbon emissions. The transition to a low-carbon future is a comprehensive societal transformation that requires collaboration among governments, industries, and individuals. Figure 2 displays the global CO2 emissions from fuel combustion by sector. What emerges, is that one key aspect of this transformation is the need to transition the heat and power sector away from fossil fuels, as it currently accounts for approximately 45% of global emissions [2].

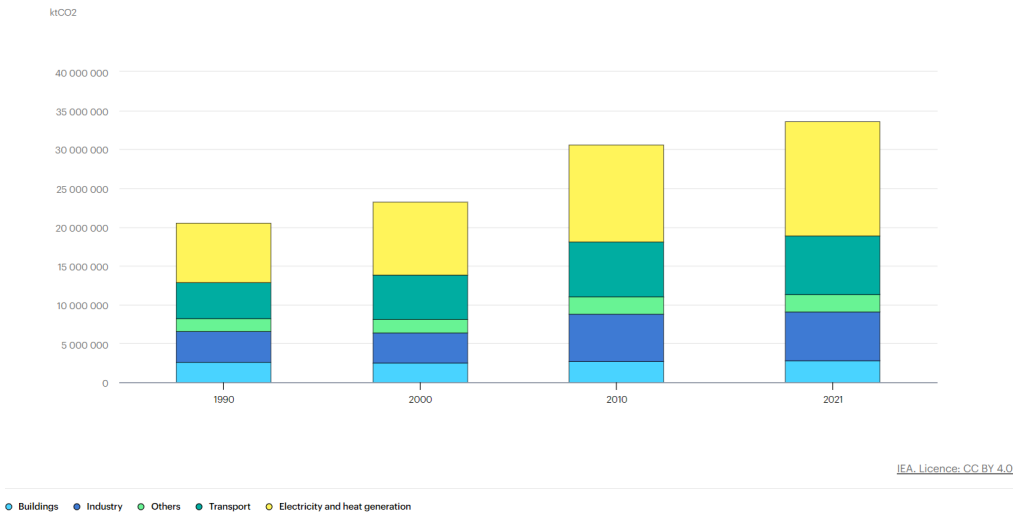

Figure 2: Global CO2 emissions from fuel combustion by sector with electricity and heat separated.

The recognition that our current energy systems are unsustainable and require a radical rethinking of how we generate, distribute, and consume energy has prompted a collective acknowledgement. Governments, industries, and communities worldwide are increasingly recognizing the importance of diversifying energy portfolios and investing in cleaner, more efficient technologies, pushing for energy transition. It refers to the global shift from conventional, fossil fuel-based energy sources to sustainable and renewable alternatives. The energy transition involves the widespread adoption of renewable energy technologies such as solar, wind, and hydropower, as well as advancements in energy storage and smart grid systems.

In this paradigm, renewable energy communities (RECs) are a promising and innovative approach to sustainable energy development. They embody the principles of local empowerment, community engagement, and environmental protection. RECs are collectives of individuals, businesses, and local organizations that collaborate to generate, consume, and often share renewable energy resources at a local level. These communities aim to promote the development and acceptance

of renewable forms of energy at the local level, energy efficiency, end-user participation in the energy market and to facilitate the provision of affordable energy to counteract energy vulnerability and poverty with environmental, economic and social benefits.

RECs represent a dynamic shift in the energy paradigm, transforming participants from consumers into proactive prosumers. They act as both producers and consumers, using renewable technologies to fulfil their energy requirements. When the energy it produces exceeds its self-consumption capacity, the prosumer sells the energy to the public grid or shares it to the members of the community. Going from consumer to prosumer is a choice in which the user finds convenience from an economic and a sustainable point of view, evolving business models related to energy management.

Establishing a Renewable Energy Community (REC) necessitates substantial investments in Renewable Energy Sources (RES), and the economic viability of these projects is a crucial factor for their successful implementation and long-term sustainability. Recognizing this, institutions are actively pushing for their widespread adoption by introducing incentivizing policies. In the Italian context, RECs are through the valorization of shared energy, an essential performance indicator for distributed self-consumption configurations, which is closely related to the synchronicity between energy generation and consumption within the community.

However, assessing this synchronicity presents significant challenges due to the inherent variability in renewable energy production, which depends on meteorological conditions, energy demand, influenced by diverse consumption patterns, and prices, strongly affected by the political and economic context. Therefore, it is crucial to consider the uncertainties in these parameters: neglecting them during a pre-feasibility study phase, only considering average profiles, may result in an unreliable assessment of the project's economic feasibility.

In this context, the objective of this thesis is to create a stochastic optimization model that focuses on identifying the optimal design for an energy production plant utilizing renewable sources, in particular photovoltaic (PV) technology. The primary goal is to maximize the Net Present Value (NPV) of the overall project. Additionally, the thesis aims to assess the feasibility of integrating an aggregate storage system within the REC.

The optimization problem is structured as a two-stage stochastic Mixed-Integer Linear Programming (MILP) model with recourse. The initial stage involves decision variables related to the sizing of both the PV and storage systems, while

the subsequent stage's variables are dedicated to optimizing energy dispatch operations. Uncertainty is introduced into the model following a probabilistic approach based on Monte Carlo simulations: different scenarios are generated, starting from historical data, in order to capture the variability of the uncertain parameters and to consistently represent their possible combinations. This comprehensive approach reinforces the strategic planning needed for effective decision-making in the establishment and sustainable operation of renewable energy communities.

The thesis is structured as follows: Chapter 1 provides an overview of European policies for sustainable development, with a focus on the definition and legislation of Energy Communities in the European and Italian context. Chapter 2 presents a state-of-the-art analysis aimed at identifying the most common approaches to the problem of optimal sizing of multi-energy systems, taking into account uncertainties in the parameters. Chapter 3 introduces the stochastic optimization framework and techniques for the goodness of fit, which are used throughout the rest of the work. Chapter 4 provides a complete description of the methodology used, including how uncertainties were modelled and how the problem was formulated. Chapter 5 evaluates the model in terms of computational times and convergence, and Chapter 6 presents the results. Finally, Chapter 7 discusses conclusions and future work.

Chapter 1

Background

Climate change poses an urgent threat, driven by human activities and carbon emissions. To combat this, a crucial shift towards renewable energy sources is underway as part of a global energy transition. Solar, wind, hydro, and geothermal energy offer sustainable alternatives, reducing reliance on fossil fuels and mitigating environmental impact. Concurrently, energy communities, comprising individuals and local entities, actively participate in and benefit from renewable projects, promoting a decentralized and inclusive energy landscape. This collective effort not only addresses climate challenges but also fosters resilience, innovation, and a more sustainable energy future for communities worldwide. In this chapter, we will give look at the strategies adopted in Europe to foster the sustainable development and how the concept of the Renewable Energy Communities has been introduces in the European context. Moreover, we will focus on RECs in the Italian regulatory framework, analyzing the normative and incentives behind them.

1.1 European Strategies for Sustainable Development

Europe has emerged as a global leader in championing sustainable development and embracing renewable energy sources to address the pressing challenges of climate change and environmental degradation. With a strong commitment to reducing carbon emissions and transitioning towards a more resilient and environmentally friendly future, European nations have embarked on ambitious initiatives and policies. The continent has witnessed a significant shift towards renewable energy, fostering innovation in wind, solar, and hydropower technologies.

The European Green Deal [3], launched in December 2019, stands as a landmark initiative defining the European Union's commitment to a sustainable and climate-neutral future. At its core, the Green Deal outlines a comprehensive strategy encompassing economic, social, and environmental dimensions to address the pressing challenges of climate change and environmental degradation. At its core is the commitment to make the EU a net-zero greenhouse gas emitter by 2050.

The Green Deal encompasses a wide-ranging set of policies and initiatives, including the Farm to Fork Strategy, Biodiversity Strategy, and Circular Economy Action Plan, all aimed at fostering a greener, more resilient economy. It promotes sustainable practices in agriculture, strives for biodiversity preservation, and seeks to decouple economic growth from resource consumption. The Green Deal addresses social and economic dimensions, ensuring that no one is left behind in the shift towards a sustainable future. With a focus on green innovation, clean energy, and circular practices, the European Green Deal is a comprehensive framework that underscores the EU's commitment to combating climate change and leading the global transition to a more sustainable and inclusive society.

In view of achieving the goals set for 2030, in 2016 the Clean Energy for All Europeans package [4], was introduced by the European Commission. It is a comprehensive initiative shaping the European Union's energy landscape. Focused on sustainability and resilience, it includes directives and regulations such as the Renewable Energy Directive and Energy Efficiency Directive. The package sets ambitious targets, requiring member states to increase the share of renewable energy, enhance energy efficiency, and improve market design. With a holistic approach, it aims to create an integrated and secure energy system, aligning with the EU's commitment to a low-carbon future and addressing climate challenges.

A cornerstone of the Clean Energy for All Europeans package, the Renewable Energy Directive (RED II) [5] is instrumental in realizing the EU's renewable energy targets. Adopted in 2018, RED II outlines binding objectives, mandating member states to collectively achieve a 32% share of renewable energy in their final energy consumption by 2030. The directive introduces measures to simplify administrative processes, enhance citizen and community involvement in renewable projects, and ensure a level playing field in the internal energy market.

RED II introduces the concept of "renewable energy communities" (RECs). These communities play a significant role in advancing the goals of a more sustainable and inclusive energy transition. RED II defines renewable energy communities as entities where members are natural persons, local authorities, or small enterprises, and their primary purpose is to provide environmental, economic, or social benefits to their

members or the local community. Key features of renewable energy communities as outlined in RED II include:

- Inclusive Membership: Members of these communities can be individuals, local governments, or small businesses. This inclusivity encourages the active participation of various stakeholders in the local energy transition.
- Community-Oriented Objectives: The primary purpose of renewable energy communities is not profit-driven but focused on delivering environmental, economic, and social advantages to their members or the local community. This ensures that the benefits derived from renewable energy projects contribute to the well-being of the community.
- Democratic Governance: RED II emphasizes democratic decision-making within these communities. Members are encouraged to actively participate in the decision-making processes, promoting transparency and community engagement in renewable energy initiatives.
- Profit Reinvestment: Any profits generated by renewable energy communities are expected to be primarily reinvested within the community or directed toward furthering the community's sustainable goals. This stipulation aims to prevent profit-maximizing activities that might compromise the community-oriented nature of these entities.
- Local Impact: RED II underscores the importance of a local dimension in renewable energy communities. The emphasis is on fostering projects that have a positive impact on the economic and social development of the local area, aligning with the EU's broader goals of regional sustainability.

The inclusion of these provisions in RED II reflects a commitment to decentralizing energy production and involving communities in the renewable energy transition. By empowering citizens, local authorities, and small businesses to actively participate in and benefit from renewable energy projects, the EU aims to build a more resilient, sustainable, and community-centric energy system. The introduction of renewable energy communities in RED II represents a significant step toward a more inclusive and people-centered approach to the clean energy transition.

1.2 Energy Communities in the Italian Regulatory Framework

The European Union is calling on its member states to comply with European directives and actively work towards achieving the set objectives for 2030. With

regards to legislation on energy communities in Italy, national regulations have incorporated the European RED II. With the DL 162/2019 ("Decreto Milleproroghe") [6], there has been the first partial transposition of RED II. It aims to initiate an experimental phase of self-consumption configurations, which will be monitored to acquire all the necessary elements for the full implementation of the European Directives. The definition of membership, member rights, and purposes remains consistent with the European directives. In Art 42bis, it is explained how legal entities formed for the establishment of energy communities and collectively acting self-consumers operate. In summary, the following conditions are imposed:

a. Participating entities generate energy for their own consumption using renewable energy plants with a total capacity not exceeding 200 kW. These plants must have come into operation after the date of entry into force of the law that converted the present decree and within sixty days following the date of entry into force of the measure that transposed the European Directive 2018/2001;

b. The energy produced is shared among the participating entities using the existing distribution network. The shared energy is the minimum amount of electricity, produced by renewable energy plants and fed into the grid, that is withdrawn by all associated end customers in each hourly period.;

c. Instantaneous self-consumption of shared energy may occur through storage systems within the designated perimeter or in the buildings or condominiums mentioned under e. The language used is clear, concise, and objective, adhering to the characteristics outlined in the assignment;

d. In the case of renewable energy communities, the consumers' withdrawal points and the feed-in points of the plants referred to in a. should be located on low-voltage electricity grids that are served by the same medium-voltage/low-voltage transformer substation at the time of the association's creation;

e. Similarly, for self-consumers of renewable energy acting collectively, they should be located in the same building or condominium.

Afterwards, the Italian Regulation Authority for Energy, Networks and Environment, ARERA, has released deliberation 318/2020/R/eel [7], which defines the tariff composition and other technical features. ARERA (the Regulatory Authority for Energy, Networks and Environment) issued Resolution 318/2020 to regulate the economic aspects of electricity subject to collective self-consumption or sharing within renewable energy communities, in accordance with Article 42-bis of Law Decree 162/19. The measure specifies that the shared energy valorisation and incentive service is provided by the Gestore Servizi Energetici (GSE) through

contact persons of groups of self-consumers of renewable energy acting collectively or of renewable energy communities. With the development of Renewable Energy Communities (RECs), a new form of participant is introduced, besides consumers and producers: the prosumers. Prosumers own the RES installations and do the physical self-consumption, instantaneously consuming the produced energy in their buildings. The surplus electricity is injected into the public grid and used to fulfil the needs of other members. According to the definition of shared electricity provided by Italian legislation, they do not contribute to the community's shared electricity.

The Ministry of Economic Development has set the incentive tariff for the remuneration of renewable source plants that are part of the experimental configurations of collective self-consumption and renewable energy communities, in accordance with Article 42-bis of Decree-Law No. 162/2019, which was converted into Law No. 8/2020. Each renewable energy plant in a renewable energy community is entitled to an incentive tariff in the form of a premium tariff for a period of 20 years. The incentive is equal to € 110/MWh for the shared electricity generated from RES installations with a nominal power lower than 200 kW. In addition, a tariff is paid for the reimbursement of network charges that have already been paid in the electricity bill for the transmission and distribution components of shared electricity. This tariff amounts to approximately € 9/MWh compared to shared energy for renewable energy communities. Community producers can also benefit from the 'Ritiro Dedicato' service by monetising the electricity they inject into the grid at the Prezzo Zonale Orario. The cost is determined by the injection hour and the market zone of the installation.

Chapter 2

State of the Art

In this chapter we will analyze the state-of-the-art for the optimal sizing of Renewable Energy Systems (RES) under uncertainty. Specifically, this literature review will delve deeper into two main topics. The first section will analyze the uncertain parameters considered in the literature and techniques for uncertainty modeling in order to comprehend how to characterize the uncertainty associated with renewable energy sources, electric power demand, and energy prices, and to explore methods for generating representative scenarios for these parameters. Secondly, we will inspect and compare the different optimization methods applied for the sizing of renewable energy systems and how energy systems have been modelled in terms of decision variables, constraints and objectives.

2.1 Uncertainty Modelling

Renewable energy sources, such as solar and wind, offer sustainable and environmentally friendly alternatives to conventional fossil fuels. Their increasing penetration raises the problem of finding new solutions to determine the right size of renewable energy systems in order to optimise their efficiency, cost-effectiveness and sustainability. However, their reliability is inherently uncertain due to their dependence on various unpredictable parameters, especially weather conditions. Solar power generation depends on the availability of sunlight, which can vary due to cloud cover, time of day and seasonal changes. Similarly, wind energy depends on wind speeds, which vary with local topography and weather patterns. These uncertainties can lead to fluctuations in energy production, making it difficult to guarantee a consistent supply of electricity. Energy demand itself can also be subject to uncertainty, depending on the behaviour of different consumers and the time of day. Most importantly, the profitability of the energy system is threatened by energy prices, which are inherently stochastic and extremely difficult to predict.

Taking these uncertainties into account is of paramount importance for the correct and comprehensive sizing of a RES.

Uncertainty modelling is a typical way of representing the stochasticity of renewable energy systems. Instead of assuming perfect knowledge of the parameters (i.e. wind speed, solar irradiation and load demand) as in a deterministic approach, random distributions are added as inputs to a stochastic optimisation approach to mimic the probabilistic characteristics of a renewable energy system. When representing the uncertainties, it is critical that the distribution dynamics of the scenarios are well captured. Both Zakaria et al. [8] and Mavromatidis et al. [9] in their works review the state-of-the-art for uncertainty models in the context of stochastic optimization for renewable energy systems. In particular, the first offers a wider view to stochastic optimization techniques for renewable energy sources applications, while the second better focuses on the characterization of the most common uncertain parameters for the sizing of RESs. Following the practise of these papers, a literature review specific for our purpose has been carried out.

The most common parameters included in the modelling of a RES are schematized in Table 2.1, while Table 2.2 shows how the stochasticity of the parameters is introduced into the optimization. The idea is to find the closest approximation to the true distributions of uncertainties. This means that we aim to infer a probability distribution for the uncertain parameters based on a given family of probability distribution functions (PDF). For this inference, Monte Carlo simulation (MCS) is one of the most widely used methods in probabilistic uncertainty modelling. Historical data on the parameters are used to determine their probability distributions, which can be used by the MCS method to learn and populate the scenario generations. The main advantages of the MCS method are its ability to systematically sample from random processes and its easy implementation. The typical use of MCS, shown in Figure 2.1, is to generate a large number of scenarios, each of which would capture the possible realisation of the underlying uncertainties.

In other cases, instead of generating a large number of scenarios, only a discrete set of possible realisations of the uncertainty factors is considered. In [10] and [18], annual data on solar radiation and wind speed are considered to generate different scenarios for the model. Mavromatidis et al. in [11] uses a probabilistic approach and applies a k-medoids clustering to select 10 representative scenarios, to which 4 additional scenarios are added to represent situations of extremely low and extremely high generation and load. Yu et al. in [17] uses the conventional stochastic scenario generation method, subdividing the load and generation distribution into three ranges, namely high, medium and low. Another way of dealing with uncertainty, as done in [12], is to define uncertainty sets, i.e. ranges

References	PV production	Solar irradiance	Temperature	Wind speed	Energy demand	Energy costs	Others
	Uncertain parameters						
Kocaman et al. [10]		✓		✓	✓		
Mavromatidis et al. [11]		✓			✓	✓	
Moret [12]						✓	✓
Rathore and Patidar [13]		✓		✓	✓		
Sharafi and ElMekkawy [13]		✓	✓	✓	✓		
Zhang et al. [15]						✓	
Zhou et al. [16]		✓		✓	✓		
Yu et al. [17]	✓		✓		✓		
Lazaroiu et al. [18]	✓						
Zheng et al. [19]		✓	✓				
Tran and Smith [20]		✓	✓				
Bashir and Sadeh [21]		✓	✓				
Reddy et al. [22]		✓	✓		✓		

Table 2.1: Uncertain parameters for sizing of renewable energy systems

References	Uncertainty modelling
Kocaman et al. [10]	Discrete scenario set
Mavromatidis et al. [11]	Representative scenarios
Moret [12]	Uncertainty range
Rathore and Patidar [13]	Probabilistic distributions
Sharafi and ElMekkawy [13]	Probabilistic distributions
Zhang et al. [15]	Agent-based modelling
Zhou et al. [16]	Probabilistic distributions
Yu et al. [17]	Representative scenarios
Lazaroiu et al. [18]	Discrete scenario set
Zheng et al. [19]	Probabilistic distributions
Tran and Smith [20]	Probabilistic distributions
Bashir and Sadeh [21]	Probabilistic distributions
Reddy et al. [22]	Probabilistic distributions

Table 2.2: Uncertainty modelling for sizing of renewable energy systems

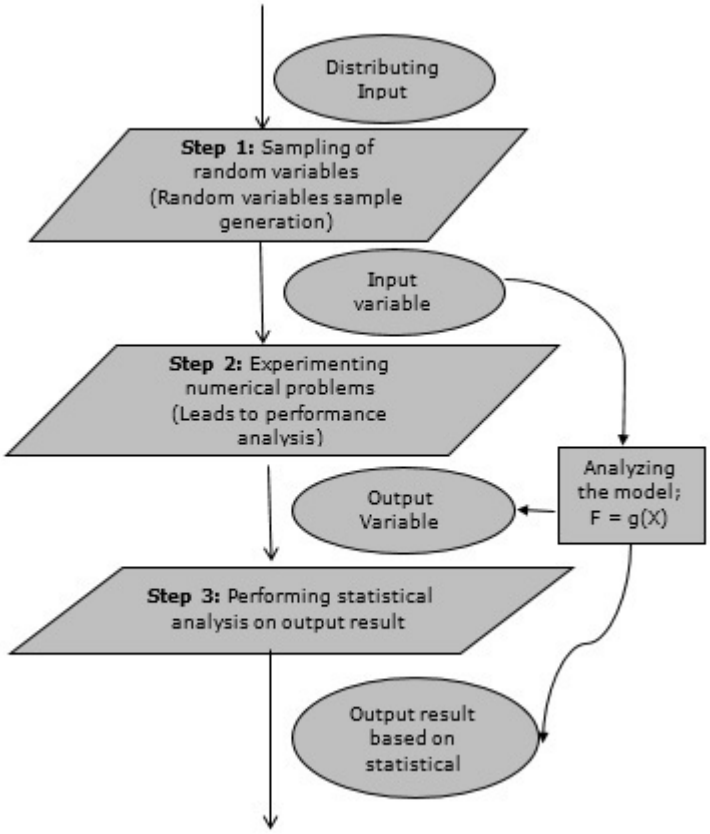

Figure 2.1: Monte Carlo simulation method

within which the uncertain parameters are known to vary. In [15], an agent-based approach is proposed to mimic the interdependent behaviour of many individuals to generate uncertain demand scenarios for a case study community. community. Finally, Zakaria et al. in [8] presents a relatively new methodology using Generative Adversarial Networks (GANs). GANs are a class of artificial intelligence algorithms used in unsupervised machine learning, specifically to generate new content or data similar to a given dataset. The concept behind GANs involves two neural networks, each with a unique role. The generator is the network that creates new data instances from random noise or a latent space representation. Its goal is to produce outputs that are indistinguishable from real data. The second network is the discriminator, which evaluates the generated data and the real data and tries to distinguish between them. Its goal is to correctly identify whether the input data comes from the generator (fake) or the real data.

Focusing on the papers using a probabilistic approach, Table 2.3 presents the

distributions typical used to fit the uncertainty of the most common parameters, solar radiance, wind speed and energy load.

Parameter	Distribution	References
Solar radiation	Normal	[14], [16], [19], [20], [21]
	Beta	[13], [22]
	Other	[11]
Wind speed	Beta	[19]
	Weibull	[13], [14], [16], [20], [21], [22]
Energy demand	Normal	[16], [22]
	Other	[11], [13], [14]

Table 2.3: Probabilitiy distributions used for uncertain parameters in the literature

The very well-know Gaussian distribution is widely adopted to represent the energy demand and solar radiation. The behavior of the latter, in addition, is also well-captured by the Beta distribution, whose PDF, shown in Figure 2.2, is given by

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{x}{x_{max}}\right)^{(\alpha-1)}\left(1-\frac{x}{x_{max}}\right), \tag{2.1}$$

where α and β are respectively the scale and shape parameter of the Beta distribution and x_{max} is the maximum value of solar radiation in the considered hour. The Beta distribution is commonly used also for wind speed, but the most popular choice in this case is by far the Weibull distribution, whose PDF, shown in Figure 2.3, is given by

$$f(x) = \frac{k}{c}\left(\frac{x}{c}\right)^{k-1}\exp-\left(\frac{x}{c}\right)^{k}, \quad x > 0, \tag{2.2}$$

where $c > 1$ and $k > 1$ are respectively the scale and shape parameters. It is important to notice that these distributions capture the possible realizations of the corresponding parameter for a given time period, most commonly of an hour.

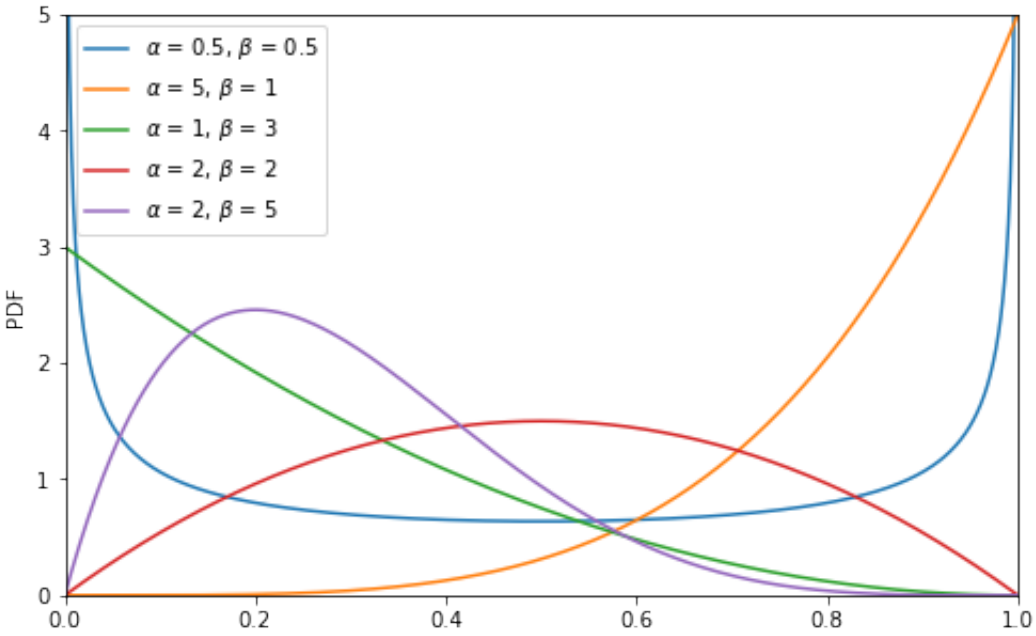

Figure 2.2: Beta distribution pdf

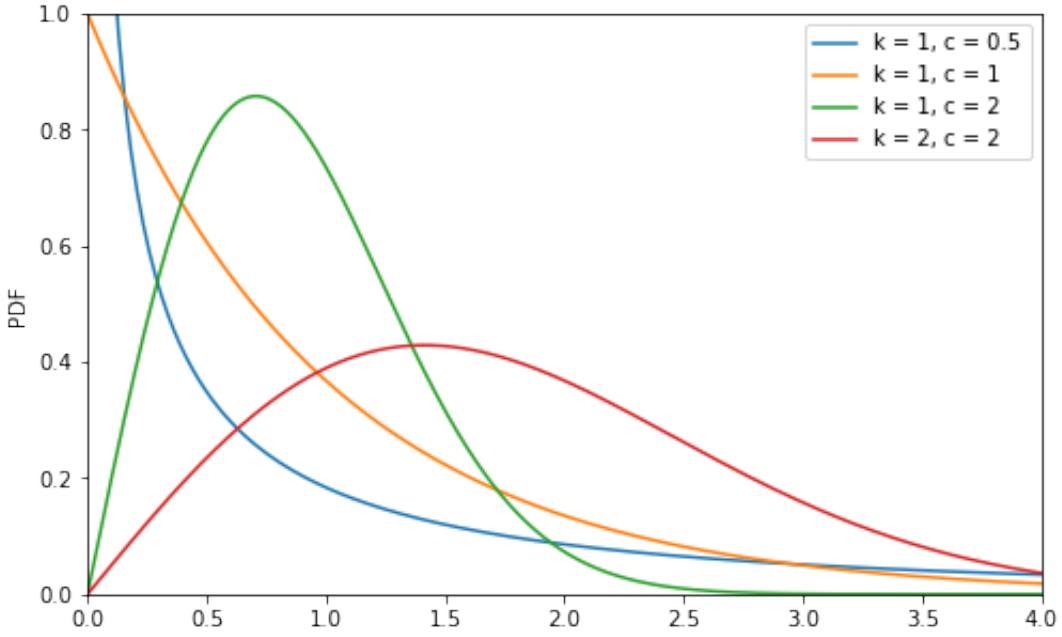

Figure 2.3: Weibull distribution pdf

2.2 Optimization methods

The following step of this literature review is the analysis of the optimization methodologies used in the context of the sizing of renewable energy systems. Also in this case, the article [8] provides some guidelines on how to evaluate the optimization process.

References	Optimization methodology
Kocaman et al. [10]	Two-stage SLP
Mavromatidis et al. [11]	Two-stage SMILP
Moret [12]	Robust optimization
Rathore and Patidar [13]	Particle Swarm Optimization
Sharafi and ElMekkawy [14]	Particle Swarm Optimization
Zhang et al. [15]	SMILP
Zhou et al. [16]	Two-stage SMILP
Yu et al. [17]	Two-stage SMILP
Lazaroiu et al. [18]	SMILP
Zheng et al. [19]	Evolutionary method
Tran and Smith [20]	Particle Swam Optimization
Bashir and Sadeh [21]	Particle Swarm Optimization
Reddy et al. [22]	SMILP

Table 2.4: Uncertainty modelling for sizing of renewable energy systems

The most commonly used method is Stochastic Programming (SP), a mathematical optimisation framework for decision making in situations of uncertainty. It deals with problems where some parameters or variables are not known with certainty, but follow probabilistic distributions. Obviously, stochastic problems can be formulated as linear programming problems, in the case of only continuous variables, or as mixed integer linear programming problems, where discrete variables are also considered in the modelling. Stochastic optimisation involves optimising decisions when the exact values of certain variables are uncertain but follow known probability distributions. Two-stage stochastic programming is a common type where decisions are made in two stages. The first stage involves decisions made before the uncertainty is revealed, and the second stage involves decisions made after the uncertain parameters have been observed. Finally, Multi-Stage Stochastic Programming extends the concept to more than two stages, allowing for multi-stage decision making in a dynamic environment with uncertain parameters. Further details of stochastic programming are given in Chapter 3. A similar approach to Stochastic Programming is Robust Optimisation, used by Moret in [12]. This

methodology is characterised by two main aspects, namely the worst-case scenario approach and uncertainty set modelling. Robust optimisation often focuses on minimising regret or loss in the worst-case scenario. Rather than seeking the best possible outcome in an ideal scenario, it seeks to minimise the impact of the worst possible realisation of uncertainty. Robust optimisation involves defining a set of possible scenarios, or uncertainty parameters, within which the solution should perform well. This set encapsulates the range of uncertainty that the decision maker wants to consider. Finally, a popular solution strategy is Particle Swarm Optimisation (PSO). It is a computational optimisation technique inspired by the social behaviour of flocking birds or schooling fish. It's a population-based metaheuristic algorithm used to find solutions to optimisation problems. The core concept of PSO is to simulate the social behaviour of individuals within a flock. In this case, potential solutions to the optimisation problem are represented as particles in a multidimensional search space. Each particle adjusts its position based on its own experience (local search) and the experience of neighbouring particles in the search space (global search).

Modelling component	References
Decision variables	
Size PV	[10], [11], [12], [13], [14], [16], [17], [21]
Size wind	[10], [12], [13], [14], [16], [21]
Size storage	[11], [12], [13], [14], [16], [17], [21]
Component selection	[11], [16]
Operational variables	[10], [11], [12], [16], [17], [18], [20]
Others	[11], [12], [14], [15], [16]
Objective	
Minimize total annualized cost	[10], [11], [12], [14], [15], [16], [21]
Minimize operational cost	[18], [20]
Minimize power loss	[13]
Minimize CO2 emissions	[14]
Others	[17]
Constraints	
Supply-demand balance equations	[10], [9], [12], [15]
Energy balance equations	[11], [13], [14], [15], [17], [18]
Operational and technical constraints	[11], [12], [13], [14], [15], [16], [17], [18], [20]
CO2 emissions	[11], [12]
Others	[10], [12], [14], [21]

Table 2.5: Variables, constraints and objectives in the literature for the multi-energy system optimization.

Table 2.5 displays the prevalent modelling components used in the revised papers, including the decision variables, objective, and constraints of the models. The

type of decision variable depends on the problem's characterization. The sizing of the photovoltaic, wind, and storage systems are recurrent variables as they are the topic of research. In papers that adopt a two-stage stochastic programming approach, the optimization of design variables is assigned to the first stage. In the second stage, operations are optimized to find the most effective demand-fulfilling dispatch strategy and use of the battery through charge and discharge. In [11] and [16] binary variables are used to represent the selection of generating components in the optimal design of the energy system, along with continuous or integer variables for sizing.

On the side of the objective, costs are of great relevance. In particular, the majority of the proposed works, optimize the total annualized cost, which comprises the initial investment cost, maintenance costs and operational costs through the lifetime of the energy system, annualized using a coefficient named Capital Recovery Factor (CRF). The CRF is the ratio of a constant annuity to the present value of receiving that annuity for a given length of time. Using a constant interest rate i and considering n annuities, the capital recovery factor is given by

$$CRF = \frac{i(1+i)^n}{(1+i)^n - 1}. \tag{2.3}$$

Instead of considering all the possible expenses, in [18] and [20] only the operational costs are minimized. In [13] the power lower is considered as objective function. Sharafi and ElMekkawy in their work utilize a multi-objective approach through the PSO. Together with the minimization of the total annualized cost, another objective is the minimization of the CO2 emission of the system. Another way to include a control of the CO2 emissions is through specific constraints, as done in [11] and [12], in which a maximum level of emissions established which cannot be exceeded. Recurring constraints are, indeed, energy balance and supply-demand equations, and technical and operational constraints, which include capacity constraints on power generators and capacity and discharge constraints for the storage systems.

Chapter 3

Theoretical Concepts

This chapter introduces fundamental theoretical concepts necessary for the development of the work. The first section analyses goodness-of-fit techniques, which are tests used to evaluate how well a distribution can represent the data. These techniques will be implemented to determine how to model the uncertain parameters of the problems, particularly the solar radiation. The following section will cover the framework of stochastic programming for modelling optimization problems that involve uncertainty. A stochastic program is an optimization problem where some or all problem parameters are uncertain, but follow known probability distributions. This framework differs from deterministic optimization, where all problem parameters are assumed to be known exactly. The aim of stochastic programming is to identify a decision that optimises a set of criteria selected by the decision maker, while also taking into account the uncertainty of the problem parameters.

3.1 Goodness of Fit techniques

This section is devoted to the presentation and discussion of goodness-of-fit techniques. By these we mean methods of examining how well a sample of data fits a given distribution as its population. These techniques usually involve formal statistical tests, and the measures of agreement or disagreement are test statistics. We will use [23] as a bibliographic reference. In the formal framework of hypothesis testing, the null hypothesis H_0 is that a given random variable X follows a specified probability law $F(x)$. Goodness-of-fit techniques used to test H_0 are based on measuring in some way the to the hypothesised distribution or, equivalently, its deviation from it. In most applications of goodness-of-fit techniques, the alternative hypothesis H_1 is composite, giving little or no information about the distribution of the data, and simply stating that H_0 is false. The main focus is on the degree of agreement of the data with the null hypothesis; in fact, it is usually hoped that

H_0 will be accepted as true. Given a sample $X_1, \cdots, X_n$, the goodness of fit test decides between the following two hypotheses about the distribution F,

$$H_0 : F \in \mathcal{F}, \quad H_1 : F \notin \mathcal{F},$$

where $\mathcal{F}$ is a family of distributions. In the case of $F = F_0$, the goodness-of-fit test becomes simple, otherwise it is known as a composite test.

The first and most popular goodness-of-fit technique is the Pearson's chi-square test, which, in synthesis, uses as a measure of contingency the sum of differences between observed and expected outcome frequencies. Suppose that n observations in a random sample from a population are classified into M mutually exclusive classes, say $E_1, \ldots, E_M$, with respective observed numbers of observations $N_1, \ldots, N_M$, and a null hypothesis gives the probability p_i that an observation falls into the i-th class, where

$$p_i = \mathbb{P}(X_j \in E_i) = \int_{E_i} dF(x). \tag{3.1}$$

Therefore, we have that the expected number of observation for class i are given by np_i. Pearson proposed that, under the circumstance of the null hypothesis being correct, as $n \to \infty$ the quadratic form

$$\chi^2 = \sum_{i=1}^{M} \frac{(N_i - np_i)^2}{np_i} \tag{3.2}$$

in limiting distribution, is the χ^2 distribution with $M - 1$ degrees of freedom. It is common to wish to test the composite hypothesis that the distribution function of the observations X_j is a member of a parametric family $\{F(\dot{|}\theta) : \theta \in \Omega\}$, where Ω is a p-dimensional parameter space. Given an estimator $\tilde{\theta_n}$ - a function of $X_1, \cdots, X_n$ - the estimated cell probabilities become

$$p_i(\tilde{\theta_n}) = \int_{E_i} dF(x|\tilde{\theta_n}), \tag{3.3}$$

and the Pearson statistics for testing the fit of the distribution $F(\dot{|}\tilde{\theta_n})$ is

$$\chi^2(\tilde{\theta_n}) = \sum_{i=1}^{M} \frac{(N_i - np_i(\tilde{\theta_n}))^2}{np_i(\tilde{\theta_n})} \tag{3.4}$$

Fisher showed that the Pearson-Fisher statistic $\chi^2(\tilde{\theta_n})$ has the distribution $\chi^2(M - p - 1)$ under the null hypothesis, no matter what θ in ω is the true value. So, for a given significance level α, we can reject the null hypothesis that the sample $X_1, \cdots, X_n$ is from the parametric family $\{F(\dot{|}\theta) : \theta \in \omega\}$, where ω, if

$$\chi^2(\tilde{\theta_n}) > \chi^2_\alpha,$$

where χ^2_α is the quantile of level $1-\alpha$ of the distribution $\chi^2(M-p-1)$.

Another class of goodness-of-fit tests is based on the empirical cumulative distribution function (ECDF). Suppose a given random sample of size n is $X_1, \cdots, X_N$ and that the distribution of X is $F(x)$. The ECDF of the sample is given by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \leq x). \tag{3.5}$$

$F_n(x)$ is a step function calculated from the data, as x increases it takes a step of height l/n at each sample observation. For x, $F_n(x)$ records the proportion of observations less than or equal to x. We can expect $F_n(x)$ to estimate $F(x)$, and indeed it is a consistent estimator of $F(x)$. The law of large numbers gives us the almost certain convergence:

$$F_n(x) \to F(x) \quad \text{a.s.}, \quad \forall x. \tag{3.6}$$

This pointwise convergence has been extended to uniform convergence. This result is known as the Glivenko-Cantelli theorem:

$$\sup_x |F_n(x) - F(x)| \to 0 \quad \text{a.s.}. \tag{3.7}$$

A statistic measuring the difference between $F_n(x)$ and $F(x)$ w ill is called an EDF statistic. The most representative statistic of this class is the Kolmogorov-Smirnov (KS) statistic, given by

$$D_n = \sup_x |F_n(x) - F(x)|. \tag{3.8}$$

3.2 Stochastic Programming with Recourse

This section introduces the theoretical concepts of stochastic optimization, with a focus on stochastic programming with recourse. This approach involves making recourse decisions, which are actions taken after uncertain events are revealed. These decisions are made in response to the actual realization of the uncertain parameters. For this section, [24] and [25] will be followed as bibliographic references.

3.2.1 Two-stage stochastic linear programming

Stochastic linear programs are linear programs where some of the problem data may be uncertain. Recourse programs allow for decisions or actions to be taken

after uncertainty is revealed. In other words, data uncertainty refers to problem data that can be represented as random variables. It is assumed that an accurate probabilistic description of the random variables is available in the form of probability distributions, densities, or probability measures. The values of the random variables are only known after the random experiment, meaning that the vector $\xi = \xi(\omega)$ is only known after the experiment. The decisions are divided into two groups:

- Before conducting the experiment, a number of first-stage decisions must be made. The initial decisions made during the first stage of the experiment are referred to as first-stage decisions.

- Subsequent decisions made after the experiment are known as second-stage decisions, and this period is referred to as the second stage.

In optimization models involving recourse, we leverage the dynamic flow of information and delineate decision stages. The basic scenario involves a two-stage decision process represented by a scenario fan, as illustrated in the Figure 3.1. In

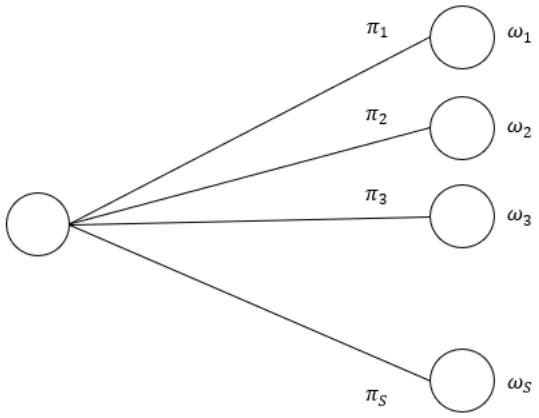

Figure 3.1: Scenario fan for two-stage stochastic programming.

the first stage, which corresponds to the root of the tree, immediate decisions are made based on the current situation. In the second stage, actions are taken based on observed random variables at the leaves of the tree. The problem to be solved

is the two-stage stochastic linear program with fixed recourse, given by

$$\begin{aligned} \min \quad & c^T x + \mathbb{E}_\omega[q(\omega)^T y(\omega)] && (3.9)\\ \text{s.t.} \quad & Ax = b, && (3.10)\\ & Wy(\omega) + T(\omega)x = h(\omega), \quad \text{a.s.} && (3.11)\\ & x, y(\omega) \geq 0. && (3.12) \end{aligned}$$

In this formulation

- The first-stage decision $x \geq 0$ must satisfy immediate constraints $Ax = b$ and lead to the immediate cost $c^T x$.
- During the second stage, a random scenario denoted as ω unfolds, introducing random data. In response to this information, a second-stage decision, represented by the recourse action $y(\omega) \geq 0$, is taken.
- Inter-stage constraints establish a connection between the second-stage decisions and the first-stage decision, as exemplified by expressions such as $Wy(\omega) + T(\omega)x = h(\omega)$.
- The cost incurred by the second-stage decision is denoted as $q(\omega)^T y(\omega)$ and is considered random from the perspective of the root node.
- The aim is to reduce the total cost by minimizing the sum of the initial cost and the anticipated second-stage cost.

If the recourse matrix W is deterministic, as above, the problem is said to be with fixed recourse. The more general case with a random recourse matrix $W(\omega)$ may present additional difficulties. The objective function of (4.1) contains a deterministic term $c^T x$ and the expectation of the second-stage objective $q(\omega)^T y(\omega)$ taken over all realizations of the random event ω . This second-stage term is the more difficult one because, for each *omega* , the value $y(\omega)$ is the solution of a linear program. To stress this fact, one sometimes uses the notion of a deterministic equivalent program. For a given realization ω, let

$$\mathcal{Q}(x) = \mathbb{E}_\xi[Q(x, \xi(\omega))] \tag{3.13}$$

be the second-stage value function, We can define the expected second-stage value function

$$Q(x, \xi(\omega)) = \min_y \{q(\xi(\omega)^T y | Wy = h(\xi(\omega)) - T(\xi(\omega))x, y \geq 0\}, \tag{3.14}$$

called recourse function, and the deterministic equivalent program (DEP)

$$\begin{aligned} \min \quad & c^T x + \mathcal{Q}(x) \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

Here, we make the presence of random variables (risk factors) $\xi(\omega)$ explicit. This formulation shows that stochastic linear programming is, in general, a non-linear programming problem. In fact the recourse function $\mathcal{Q}(x)$ is nonlinear, in general, and presents some complications:

- It is an expectation with respect to the joint distribution of $\xi(\omega)$; hence, it is a multidimensional integral, if random variables are continuous.
- It is a multidimensional integral of a function that we do not really know, as it is implicitly defined by an optimization problem.

Luckily, in many cases of practical interest, we can prove interesting properties of the recourse function, most notably convexity. In some cases, $\mathcal{Q}(x)$ is differentiable; in other cases it is polyhedral. Another advantage of this view is that it makes the transition to multistage stochastic programming and stochastic dynamic programming more natural. We make a decision x here and now, but we cannot only focus on its immediate cost: The future cost-to-go must be accounted for, which is the purpose of the recourse function.

3.2.2 Multi-stage stochastic linear programming

Multi-stage stochastic programming formulations are a natural generalization of two-stage models. The multistage stochastic linear program with fixed recourse can be defined as follows:

$$
\begin{aligned}
\min \quad & c_0^T x_0 + \mathbb{E}_{\xi_1}[\min c_1^T(\omega) x_1(\omega_1) + \cdots + \mathbb{E}_{\xi_H}[\min c_H^T(\omega) x_H(\omega_H)] \cdots] \\
\text{s.t.} \quad & W_0 x_0 = h_1, \\
& T_0(\omega_1) x_0 + W_1 x_1(\omega_1) = h_0(\omega) \\
& \vdots \\
& T_{H-1}(\omega_H) x_{H-1}(\omega_{H-1}) + W_H x_H(\omega_H) = h_H(\omega), \\
& x_1, x_t(\omega_t) \geq 0, \quad t = 1, \ldots, H.
\end{aligned}
$$

To understand this model, we simply need to nest resource functions that correspond to decision stages. The sequence of decision stages can be comprehended by following the logical structure of the model:

- When the first time period starts, at time $t = 0$, the decision vector x_0 is selected; this decision has a deterministic immediate cost $c_0^T x_0$ and must satisfy the constraint

$$W_0 x_0 = h_0.$$

- At the beginning of the second time period, at time $t = 1$, the random vector

$$(T_0, h_1, c_1)$$

is observed and after, using this information, the decision x_1 is taken; this decision has an immediate cost $c_1^T x_1$ and must satisfy the constraint

$$T_0(\omega_1)x_0 + W_1 x_1(\omega_1) = h_0(\omega).$$

We notice that this information is not known at time $t = 0$, but only at time $t = 1$; we will make the new decision on the basis on the of these random variables, also considering the previous decision.

- The above procedure is repeated for all of the time periods up to H_1, where H is our planning horizon.

- Lastly, at the start of the last time period H, the random data (T_{H-1}, h_H, c_H) is observed; then, on the basis of this information, we make decision x_H, which has an immediate cost $c_H^T x_H$ and must satisfy the constraint

$$T_{H-1}(\omega_H)x_{H-1} + W_H x_H(\omega_H) = h_0(\omega).$$

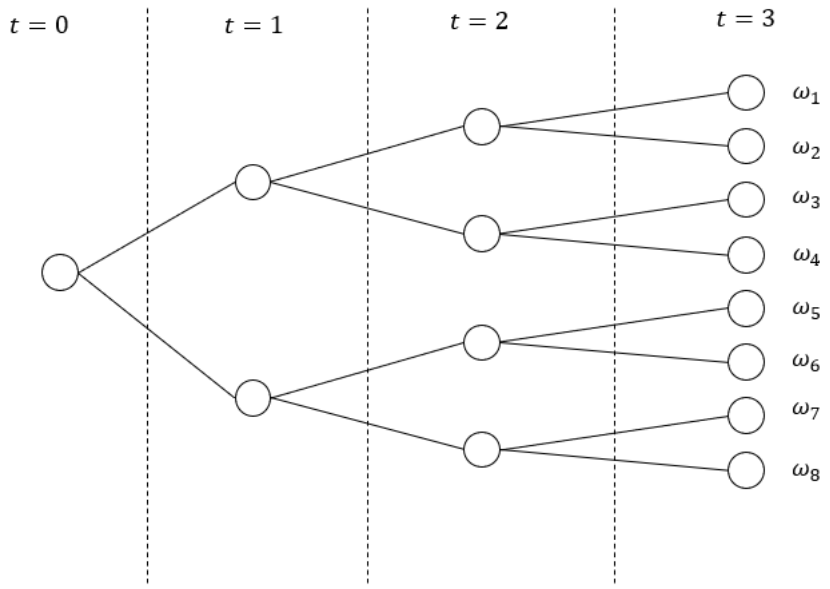

Figure 3.2: Scenario tree for multi-stage stochastic programming.

3.2.3 Sample Average Approximation

In this section, the Sample Average Approximation (SAA) for stochastic problems is presented. SAA is a technique used in stochastic programming to handle uncertainty by approximating the expected value of an objective function or constraints using a sample of scenarios. The general idea is to replace the true expected value in the objective function with a sample average over a finite set of scenarios. This allows for solving a deterministic optimization problem that approximates the original stochastic problem. The sample average is calculated based on the outcomes of multiple scenarios randomly sampled from the distribution of the uncertain parameters.

Suppose that we have a sample $\xi^1, \ldots, \xi^N$ of N realizations of the random vector ξ. This random sample can be viewed as historical data of N observations of ξ, or it can be generated in the computer by Monte Carlo simulations methods. The most direct approach to the two-stage stochastic program is to replace the recourse function, $\mathcal{Q}(x)$, by using the estimate

$$\mathcal{Q}^N(x) = \sum_{k=1}^{N} \frac{Q(x, \xi^N)}{N}. \tag{3.15}$$

This then yields the sample average approximation problem for the stochastic linear program as

$$\begin{aligned}
\min \quad & c^T x + \frac{1}{N} \sum_{k=1}^{N} q_k^T y_k && (3.16)\\
\text{s.t.} \quad & Ax = b, && (3.17)\\
& T_k x + W y_k = h_k, \quad \forall k = 1 \ldots, N && (3.18)\\
& x, y_k \geq 0. && (3.19)
\end{aligned}$$

Chapter 4

Methodology

The methodology chosen for the developed solution utilises a two-stage stochastic programming framework solved through the sample average approximation method. The main objective is to optimise the size of the energy systems, specifically the number of photovoltaic units and batteries, while taking into account the uncertainty of the energy demand, renewable energy generation, energy load, and prices. The second stage of the stochastic optimization process involves generating numerous scenarios, each comprising different realizations of the uncertain parameters sampled hourly for a year. This chapter provides a detailed analysis of the steps involved in scenario generation and concludes with the formulation of the optimization model.

4.1 Model Components

This section provides a detailed discussion of the input variables and parameters of the problem, as well as the mathematical modelling of the system. The optimization phase is solved using the sample average approximation method, which is based on Monte Carlo simulation to approximate the second stage cost of the stochastic optimization model. The Monte Carlo simulation method generates scenarios of possible years. Each scenario consists of a year of hourly profiles of photovoltaic production, energy load, and energy prices. Therefore, we will analyze how to model these uncertain parameters and integrate them into the optimization process. The case study for this work is a hypothetical energy community in the city of Taranto (Puglia, Italy). We present the results of the analysis for this case study, but the methodology can be generalized for any location.

4.1.1 Photovoltaic generation

The common approach in the literature, as seen in Chapter 2, is to model the photovoltaic generation starting from data of solar radiation and temperature, as well as technical specifications of the chosen technology. However, this approach may present difficulties, particularly in defining a correlation between solar radiance and temperature to accurately predict PV output. Therefore, data on photovoltaic energy production will be used to construct a probability model for the generation side. This data can be obtained from PVGIS, an online tool developed by the European Commission's Joint Research Centre, which provides information and tools to evaluate the solar energy potential at a specific location.

Using a probabilistic approach, a probability distribution is assigned to each hour or time band of a specific month or season to represent solar radiation and temperature in that time interval. Hourly profiles are then generated from these distributions. In this work, a random variable of PV generation will be defined for each hour of each month. The aim is to create an hourly profile of PV generation over one year that accurately reflects real data. To reduce the number of parameters to estimate and data to process, and consequently decrease computational times, we will only consider a collection of 12 PV generation profiles per year, one for each month.

After establishing the methodology, the next step is to select the most suitable probability distributions for PV generation. Three potential distributions have been examined, including the commonly used Normal and Beta distributions. Histograms of the frequencies of PV generation for three specific hours (10:00, 12:00, and 14:00) in January and July are presented in Figures 4.1 and 4.2. A qualitative analysis of the distribution of PV generation values within an hour of a month reveals that the data does not fit the aforementioned uni-modal distributions. Instead, it exhibits multiple peaks. Therefore, the third distribution in the comparison will be the Gaussian Mixture distribution. Its pdf is obtained as a convex combination of the pdfs of a given number of Gaussian variables with a given mean and covariance matrix. The coefficients of the combinations are referred to as the weights of the components.

Figure 4.3 provides a qualitative overview of how the data fits the candidate distribution. It is evident that the Normal distribution is unable to accurately identify the main peak in the data and appears to be quite flat. In contrast, the Beta distribution is better suited to the data, although it tends to overemphasize extreme upper values. Of the three models, the Gaussian Mixture model appears to be the best fit for the data, although it may be prone to overfitting and this needs to be verified.

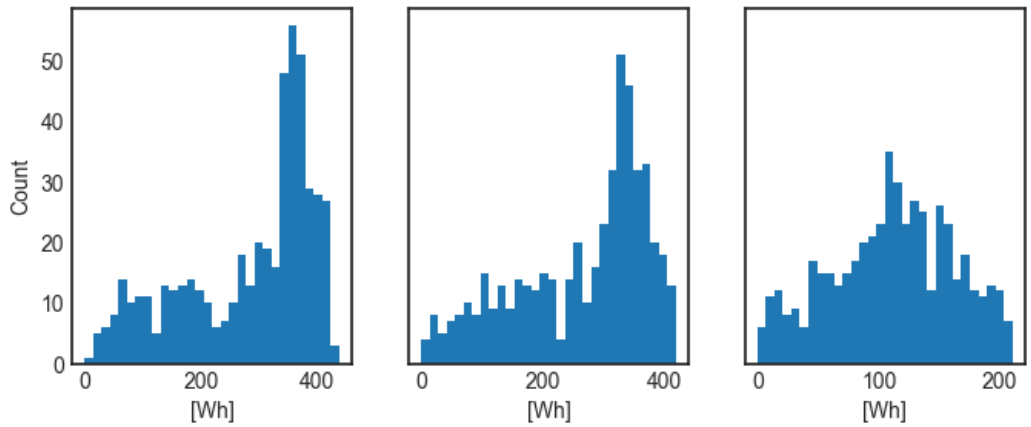

Figure 4.1: Frequencies of PV generation in January.

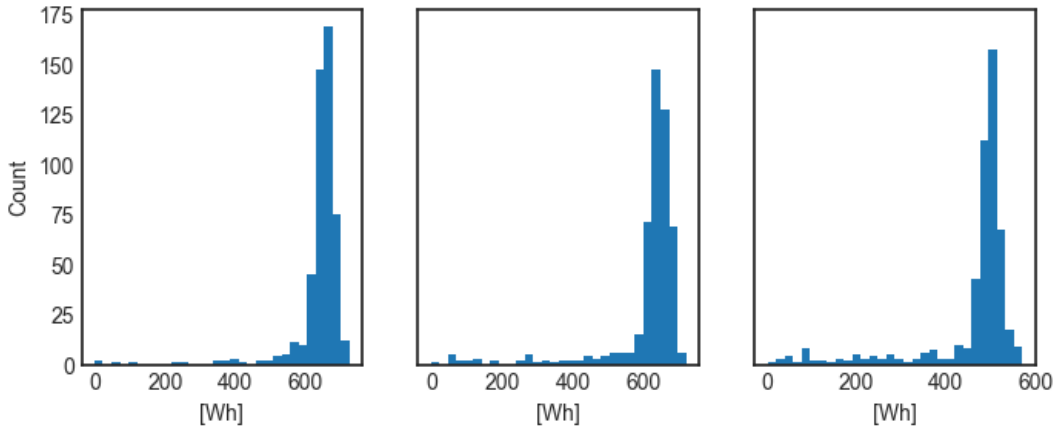

Figure 4.2: Frequencies of PV generation in July.

The choice among the three alternatives was made by performing the Kolmogorov-Smirnov test and comparing the values of the KS statistic and p-values of the models. Table 4.1 presents the results of tests conducted for the distributions of the hours in January. Based on this evaluation, we conclude that the Gaussian Mixture model is the most suitable for our application. Therefore, it will be implemented to represent the PV generation.

One subtle issue that requires further evaluation is that the Gaussian Mixture distribution has an unbounded probability density function. Therefore, although with low probability, sampling from it could result in negative values of PV generation. The first plot in Figure 4.4 displays for each month and each hour the

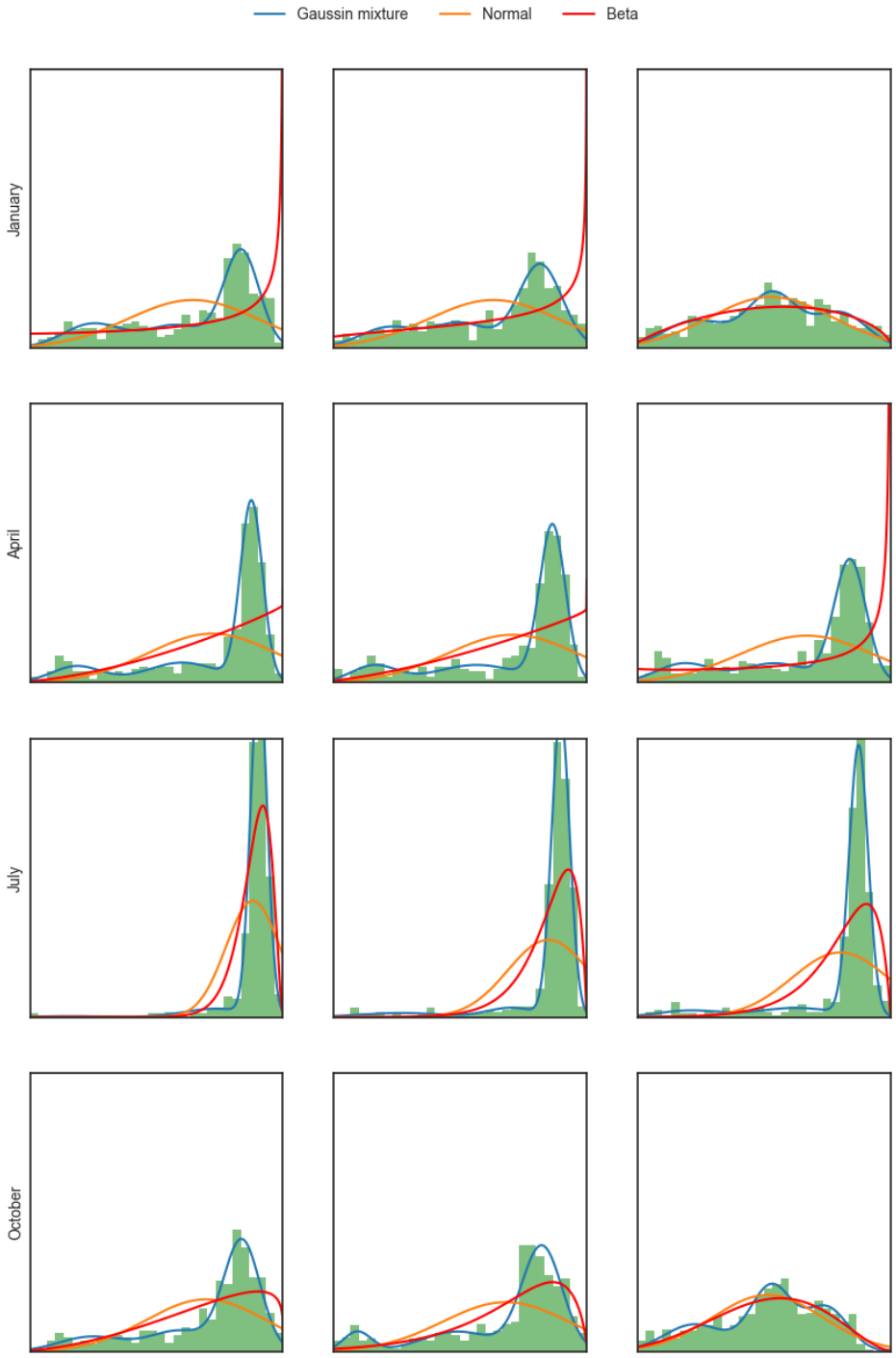

Figure 4.3: Fitting of distribution to PV generation frequencies.

	Gaussian Mixture		Normal		Beta	
Hour	D	p-value	D	p-value	D	p-value
07:00	0.60484	0.324675	0.066532	0.222422	0.969758	0
08:00	0.050403	0.554805	0.135081	0.000230	0.991935	0
09:00	0.062500	0.287476	0.183468	0	0.997379	0
10:00	0.042339	0.766177	0.215726	0	0.997984	0
11:00	0.044355	0.714273	0.165323	0.000002	0.975806	0
12:00	0.030242	0.977364	0.118952	0.001780	0.991935	0
13:00	0.060484	0.324675	0.135081	0.000204	0.965726	0
14:00	0.074597	0.126570	0.064516	0.253413	0.991935	0
15:00	0.064516	0.253414	0.125000	0.000851	0.953629	0

Table 4.1: Kolmogor-Smirnov test results for January.

probability for the fitted Gaussian Mixture variable to be negative, while in the second in Figure 4.5 the average over all the negative values over 1000 samples is computed. What emerges is that the probability of negative values is negligible in most of the cases, except for twilight hours, when it can over 40%. On the other side, during these periods, the average over the negative value is extremely low. These considerations imply that no such an error will be made by handling by simply casting to zero the sampled negative values.

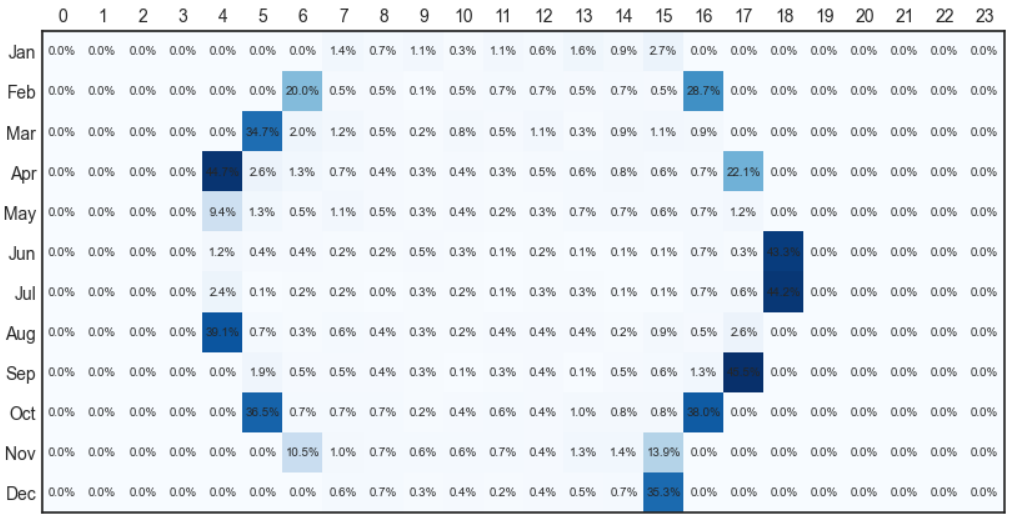

Figure 4.4: Percentage of sampled negative values of PV production for each month and each hour.

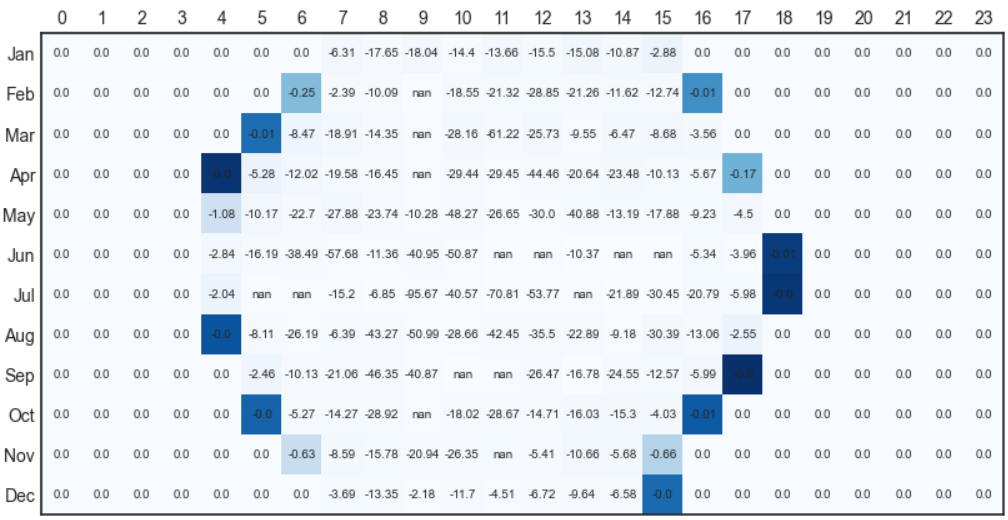

Figure 4.5: Minimum sampled negative values of PV production for each month and each hour.

To summarise, the methodology for generating PV production scenarios, consisting of consecutive hourly profiles from each month, will be to sample from the fitted Gaussian Mixture distributions in sequence. Figure 4.6 shows examples of hourly profiles from January and July.

The production of photovoltaic energy by panels is affected by both meteorological conditions and the azimuth and slope of the panels. Figure 4.7 demonstrates the effect of different slopes and azimuth values on energy production. Various parameter configurations result in scaled and shifted production profiles, which may be more advantageous depending on the periods of higher energy demand from users. For instance, panels facing east generate profiles that are shifted towards morning hours, while those facing west generate profiles that are shifted towards afternoon hours.

Considering these factors, it is important to parameterise the slope and azimuth for each building. For the sake of simplicity, we will assume that each building has a single pitch, characterised by an azimuth angle $\alpha \in [90,270]$ - where 90° corresponds to east and 180° to south - and a slope $\beta \in [10,40]$. In addition, a parameter representing surface will be assigned to each building. The surface is strictly related to the maximum number of installable panels, thus will set a maximum-capacity-type constraint to the PV sizing.

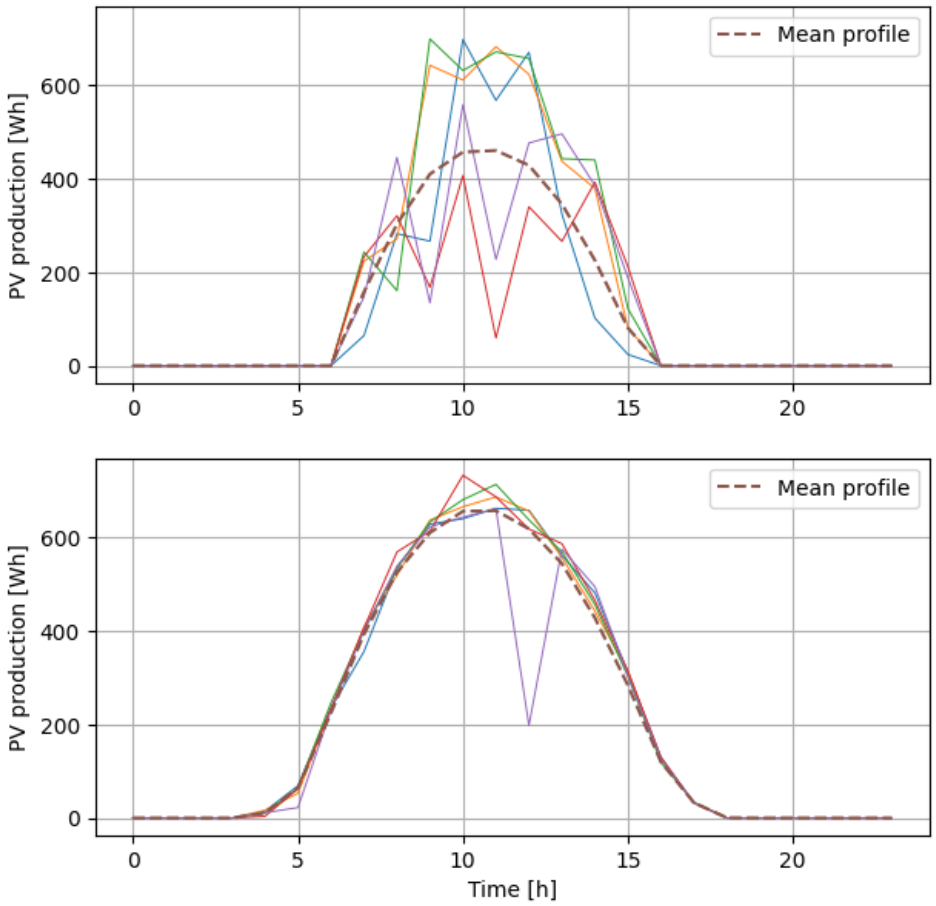

Figure 4.6: Examples of generated PV profiles for January (top) and July (bottom).

4.1.2 Energy community with storage

The introduction of a storage system into the REC will now be discussed. Two alternative schemes have been identified: individual storage, in which each building has its own storage system for individual self-consumption, and a centralized configuration, in which the community has a shared storage system that can be charged by the energetic surplus of every building and can be used to cover the demand of every member of the community. The centralized configuration was chosen.

Each battery is characterized by a maximum capacity C_{max}, i.e., the maximum level of energy that can be stored in it, an efficiency coefficient η, a measure of the amount of energy that can be delivered by the battery with respect to that stored,

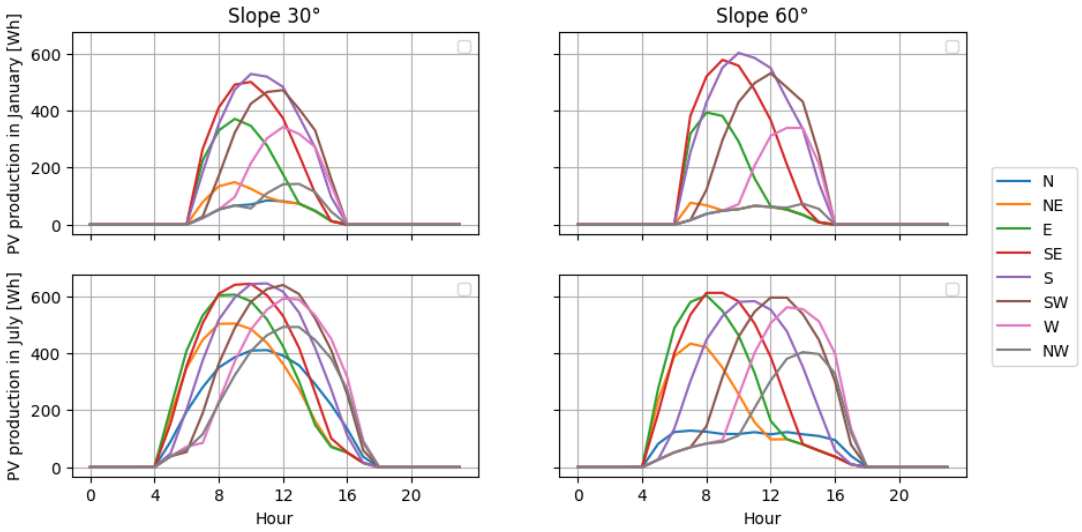

Figure 4.7: PV generation for different values of slope and azimuth.

and for each time interval a level of charge. In the case of individual storage, buildings can use their surplus energy to charge only their own batteries. In the second case, batteries can be charged by any producer or prosumer in the community. The process of charging a battery is subject to certain constraints. During the charging process, the level of charge cannot exceed the maximum capacity of the battery. Additionally, a battery cannot be fully charged in a single session; instead, the amount of energy used for charging cannot exceed a fraction of the maximum capacity. Similar constraints are applied to the discharge of the batteries. The discharge cannot exceed the current level of charge and cannot be more than a fraction of the maximum capacity.

Another important constraint is that each battery cannot charge and discharge simultaneously in a single time slot, which can be expressed through a big-M constraint. Imposing this type of constraint for each installed battery might lead to a considerable computational effort, since binary variables would have to be introduce into the formulation of the problem. Therefore, we will restate this constraint in an aggregate way, considering, for a given time instant, the number of fully or partially charged batteries, say $N_{charged}$. In this setting, the discharge cannot exceed the present level of charge of the storage system, while it possible to charge only completely empty battery. If we let x_S be the number of installed batteries, than the charge has to be less than $(x_S - N_{charged})C_{max}$. This means that the one partially charged battery will be only used for discharging, as shown in Figure 4.8.

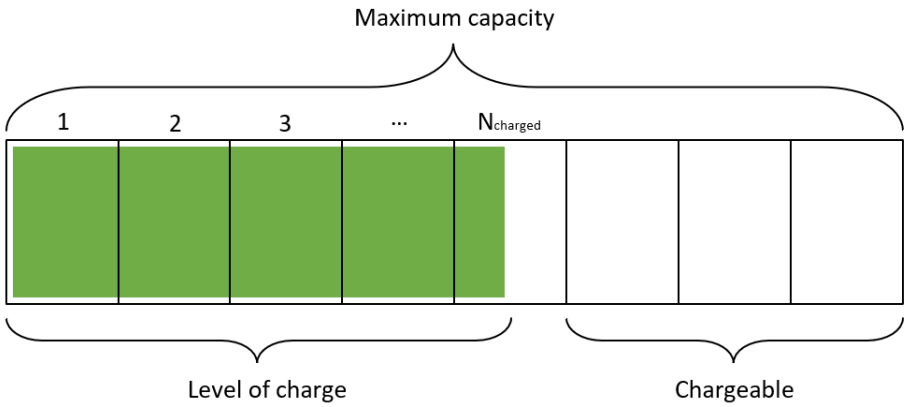

Figure 4.8: Storage system scheme.

The modelling for the actual energy communities designed to accomplish two requirements. On the prosumer side, it necessary to have access to individual profiles of generation and load and information about physical self-consumption for each hour, while from the community perspective, we want to keep trace of the discharging and charging fluxes of the battery, the energy exchanges with the grid and the value of shared energy each hour.

Specifically, we assume that each member of the community, at each hour, firstly physically self-consumes. The self-consumption can lead to either a surplus or a deficit of energy, which are both evaluated in an aggregate way, from the perspective of the REC. As matter of fact, a surplus of the community will be defined, which is given by the sum of the surplus of the individuals. This common surplus can be either used to charge the batteries or it can be imported into the grid. Similarly, for the deficit of the REC, which can by covered either by the storage system or through a withdrawal from the grid. This approach matched with the overall objective function - which will be later discussed - which evaluates the economic performance of the REC, in an aggregate way.

4.1.3 Energy demand

Residential users represent a relevant component within a REC. A correct characterization of the energy demand of the different residential buildings is, thus, of paramount importance. Compared to other classes, residential of building exhibit extremely variable patterns of electrical consumption, due to different social classes, composition and behaviors of households. Schiera et al. in [26] performed a social

clustering of the households in nine social groups and evaluated they frequencies within the district of San Salvario in Turin, in order to better comprehend the diffusion of PV technology in the urban context. Nine social groups have been identified: Ruling class (RC), Silver Pensioners (SP), Clerks' households (CH), Young blue-collars (YB), Retired blue-collars' households (RB), Lonely old ladies and young unemployed (LY), Traditional provincial households (TP), Low-income Italian households (LI), Low-income households with foreigners (LF). Moreover, for each cluster typical electricity load profiles throughout a year have been obtained, shown in Figure 4.9. This particular distribution of social groups has a strong

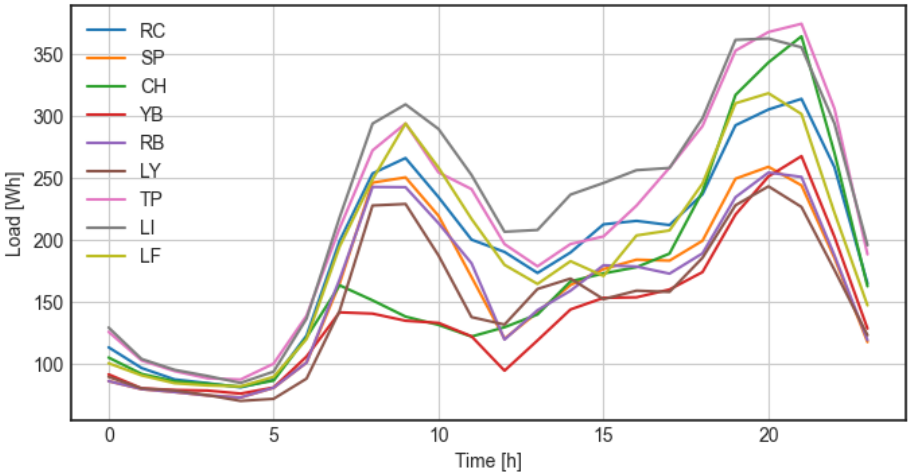

Figure 4.9: Average hourly energy load for the consumers' classes over a year.

influence on the electricity load profiles. In fact, groups CH and YB have a typical single night peak demand. As a result, we expect that the majority of households will not be able to make good use of PV production due to the significant time mismatch between load and PV production profiles. In this work we want to characterize the residential loads in a similar fashion, applying that approach with the data of the location under analysis.

The applied methodology is to obtain for each social group a standardized per capita load profile, to be scaled based on the average per capita energy consumption and a simulated number of members of the household. In this way, we can simulate a load profile for each residential building, knowing the number of housing units in it and the frequencies of social groups for the considered area. The data of per capita consumption are collected from the statistical publications of Terna [27].

The standardized profiles are easily obtained dividing the given profile by the total annual energy consumption - computed by taking the sum of hourly-resolution

data of power consumption of the profile over a year - and dividing once more by the average number of household members for the social groups. For a better representation of load profiles, we assume that the number of members for each household to be random. We modelled this number, through a triangular distribution $\mathcal{T}(min, mode, max)$, differentiating the parameters according to the social groups. The used parameters are represented in Table 4.2.

	RC	SP	CH	YB	RB	LY	TP	LI	LF
min	2	2	2	2	1	1	3	3	2
mode	2.5	2.2	2.7	2.1	1.8	1.5	4.3	4.3	2.6
max	3	3	3	3	2	2	6	6	3

Table 4.2: Parameters for the triangular distribution representing the number of members of the social classes.

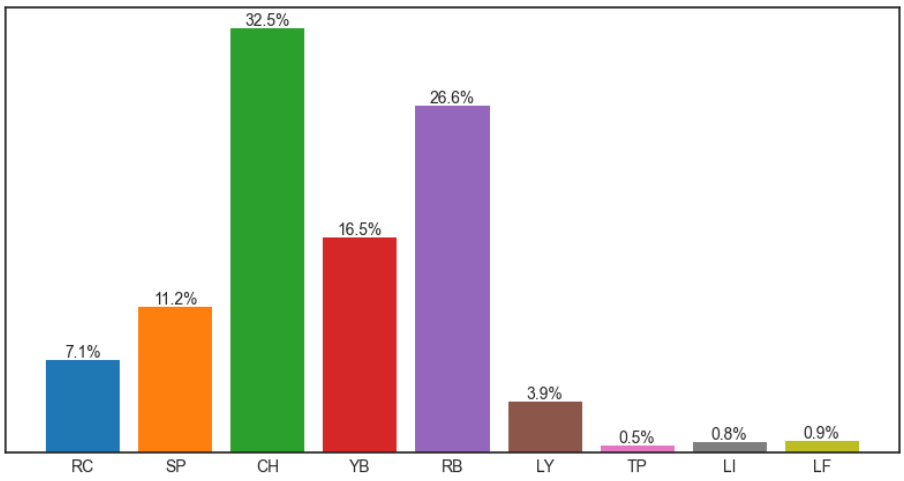

Figure 4.10: Frequencies of the social classes.

Finally, we simulated the total energy demand of a residential buildings, supposing it to be composed by 10 apartments. For each apartment, a social group has been sampled according to the specified frequencies in Figure 4.10 and the number of members of the household. Accordingly to the data scheme of PV generation, the profiles have been grouped by hours and months.

Besides residential buildings, a REC is commonly composed businesses or public entities. In particular, we will consider four other classes of buildings, which are schools, restaurants, offices and stores. We model their load on the basis of

estimated mean monthly load profiles produced by ARERA [28]. These profiles are shown in Figure 4.11. Each type of building clearly exhibit a different pattern in the

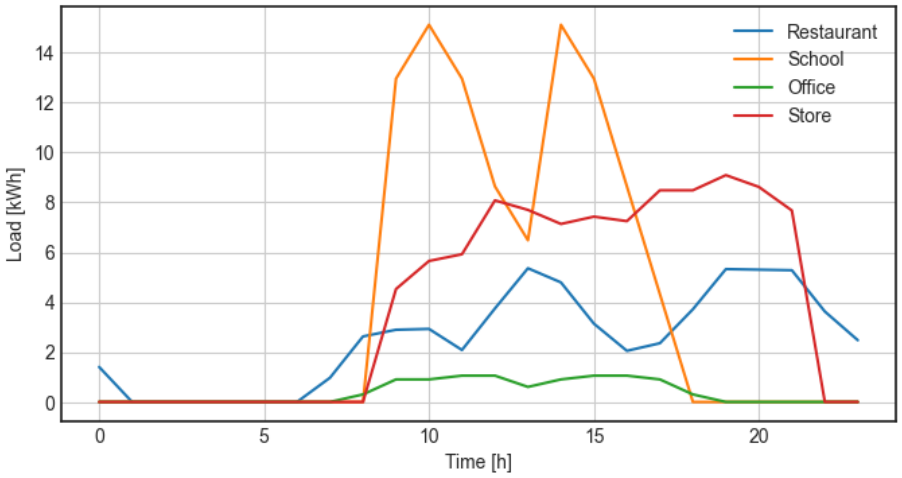

Figure 4.11: Average energy load profiles for non-residential buildings.

load profiles. As an example, schools exhibit a double-peaked profile, especially in the morning and afternoon hours, while offices have a pattern similar to residential buildings. Restaurants, instead have a consumption more shifted towards the night.

In order to increase the variability of the load modelling, we introduced some Gaussian noise as done in [16]. In particular, it is assumed that probability distributions of energy demands at each sampling time obey normal distribution $\mathcal{N}(\mu, \sigma^2)$, where μ represents the value of each point on the demand profile and 1.96σ equals 20% of μ because a 95% of confidence interval of normal distribution is corresponding to 1.96σ.

4.1.4 Energy prices

The significance of energy prices in influencing the outcomes of our optimization model cannot be overstated, especially considering the inherently economic nature of the objective function. Energy prices act as a critical determinant in shaping the cost structure of our model, and thus, their accurate representation is paramount for a comprehensive understanding of the underlying uncertainties.

In our modeling framework, we make a crucial distinction between two types of energy prices, contingent upon the nature of the interaction with the grid. The valuation of energy imported into the grid is intricately linked to the Prezzo Unico Nazionale (PUN), a dynamic variable that undergoes hourly fluctuations. This

nuanced approach ensures that the model is responsive to the real-time variations in energy import costs, capturing the temporal intricacies that contribute to the overall energy landscape.

The Prezzo Unico Nazionale (PUN) is an Italian term that translates to "Unique National Price." It refers to the hourly wholesale price of electricity in Italy. The PUN serves as a crucial benchmark for determining the cost of electricity at any given hour and is used to valorize or assess the value of energy being imported into the grid.

The PUN is a dynamic and variable metric, changing on an hourly basis, reflecting the supply and demand dynamics in the Italian electricity market. It is influenced by various factors such as energy production, consumption patterns, market regulations, and external economic and political influences. The continuous fluctuations in the PUN allow for a more granular and responsive pricing mechanism, aligning with the real-time conditions of the electricity market.

In energy models and optimization strategies, understanding and incorporating the PUN is vital. The valorization of energy imported into the grid is directly linked to the current PUN, ensuring that the model accounts for the varying costs associated with electricity procurement at different times of the day. This dynamic pricing approach helps in optimizing decisions related to energy consumption, production, and storage, considering the economic implications associated with the ever-changing PUN.

Conversely, the price associated with withdrawing energy from the grid is specifically determined by the billing structure. This fixed cost for energy withdrawal forms an essential component of our modeling strategy, providing a stable reference point within the dynamic context of energy transactions. By distinguishing between these two types of energy prices, our model is equipped to handle the multifaceted nature of energy interactions, acknowledging the distinct dynamics that govern the inflow and outflow of energy within the grid.

To provide a comprehensive understanding, Figure 4.12 illustrates the average annual PUN spanning the years 2005 to 2023. The trend, until 2020, exhibits a gradual and uniform decrease. However, the trajectory takes a distinctive turn in recent years, marked by fluctuations attributed to macro-economic and political factors. Recognizing the anomalous nature of these recent years, we have identified them as outliers and consequently omitted them from subsequent analyses.

To enhance understanding of the trend, an additive decomposition of the hourly

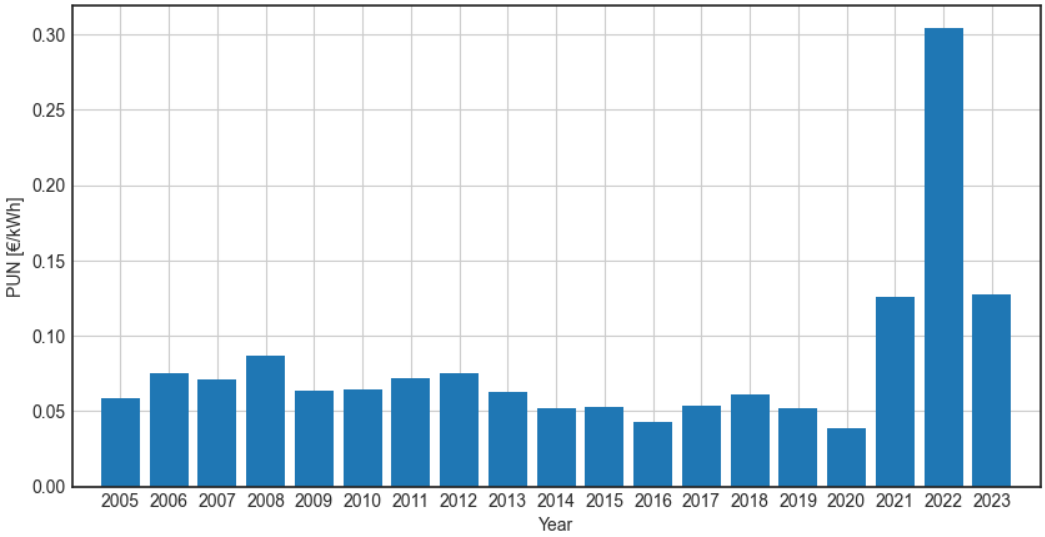

Figure 4.12: Mean annual PUN from 2005 to 2023.

data for the PUN was performed. The time series, $Y(t)$, was decomposed into three components: the trend, $T(t)$, the seasonality, $S(t)$, and the residuals, $E(t)$. A trend is present when there is a long-term increase or decrease in the data, and The presence of seasonality in a time series is indicated by a seasonal pattern, which is caused by seasonal factors such as the time of the year or the day of the week. Seasonality always occurs at a fixed and known frequency. The residual accounts for the noise present in the data. As step for the seasonality we consider a trimester. The additive decomposition reconstructs the original time series as follows

$$Y(t) = T(t) + S(t) + E(t). \tag{4.1}$$

These results are shown in Figure 4.13. As we mentioned earlier, removing the seasonality effect and the residuals, it is clear evident the decreasing trend of the PUN. We utilized this trend to define some informative levels for the value of the PUN. The core idea is, in fact, to set up three indicative values for the average annual PUN, namely low, average and high, based on the trend. In order to do so, the range of all the values of the trend have been divided into three uniform bands, and for each band the mean value for the PUN has been assigned. Having defined the value for the low, middle and high scenarios, the next step is to associate to them a probability for their realization. For this purpose, we evaluated the number of occurrences of each band compared to the total number of observations, taking their ratio to obtain the probability. To sum up, the low, average and high scenarios are specified in Table 4.3.

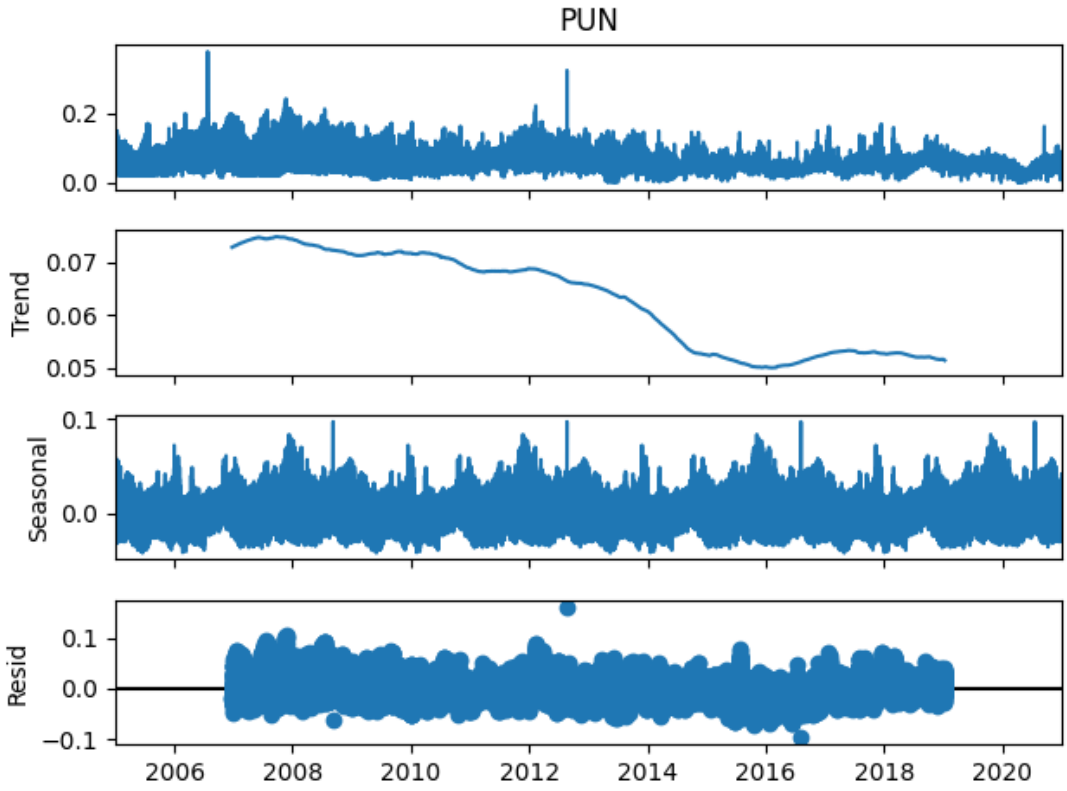

Figure 4.13: Additive time decomposition for the PUN.

Scenario	Value [€/kWh]	Probability
Low	0.0524	0.3975
Average	0.0622	0.1338
High	0.0717	0.4687

Table 4.3: Low, average and high scenario values for the average annual PUN and respective probabilities.

The next step for the PUN modelling is to reconstruct an hourly profile starting from the average annual value. This reconstruction has been differentiated for each of the three scenarios. Based on the belonging scenario, a weight has been assigned to each hour, computed as the average ratio between the value of the PUN in that hour over the average annual PUN. Through these weights we then reconstructed a average hourly profile by multiplying them by one of the average annual values found before. Finally, to increase the variability between the generated profiles a

random Gaussian noise has been added, based on the historical data.

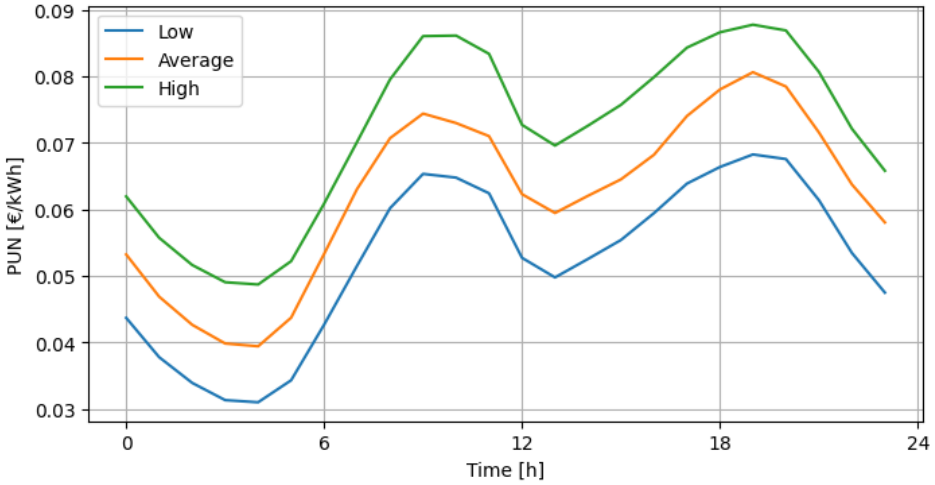

Figure 4.14: PUN average profile reconstruction starting from its annual average value.

The energy import from the electrical grid are regulated through the bill prices, which we assume, for simplicity to be equal for each member of the community and fixed throughout the whole year. Figure shows the trend of the average bill prices, evaluated quarterly, and how this price has been partition among the different voices that compose it, namely energy cost, system charges, transport and meter management costs and taxes. Also in this case, data from 2022 and 2023 has been excluded by the analysis, since considered extreme cases, not representative of the general trend. Applying the same methodology used for the PUN, we obtained three scenarios of bill prices, summarized in Table 4.4.

Scenario	Value [€/kWh]	Probability
Low	0.1414	0.2031
Average	0.1719	0.3750
High	0.2024	0.4219

Table 4.4: Low, average and high scenario values for the average bill price and respective probabilities.

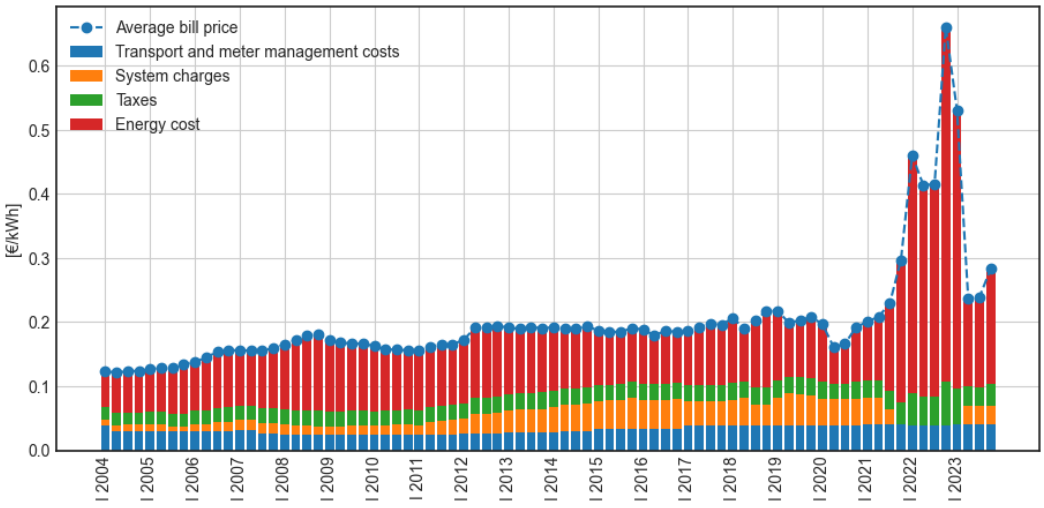

Figure 4.15: Average quarterly bill prices from 2004 to 2023.

4.2 Model Formulation

In this section, we will discuss about the mathematical formulation for the sizing of the REC, analyzing the variables, constraints and objective function of the model. The formulation is based on the two-stage stochastic programming framework. Therefore, first and second stage variables will be distinguished. The first stage aims to determine the optimal sizing of the REC, while the second stage accounts for the energy dispatch operations of the community.

4.2.1 Parameters, variables and objective

Indices and Sets

- $\mathcal{N} = \{1, \ldots, N\}$: set of the building of the problem
- $\mathcal{H} = \{0, \ldots, H-1\}$: set of the hours of the problem
- $\mathcal{S} = \{1, \ldots, S\}$: scenario set

First stage decision variables

These decision variables represent the decision "here-and-now" of the optimization problem and determine the size of the energy system:

- $x_{PV}^{n} \in \mathbb{N}$: number of photovoltaic units to install for building n

- $x_S \in \mathbb{N}$: number of storage units to install

Second stage decision variables
These decision variables represent the decision "wait-and-see" of the optimization problem. They represent the optimal operational decision to apply to optimize the objective function after the realization of the risk factors:

- $P_{grid}^{h,s} \in \mathbb{R}^+$: energy withdrawn from the electrical grid at hour h in the scenario s $[kWh]$
- $E_{+}^{h,s} \in \mathbb{R}^+$: discharged energy provided by the storage system at hour h in the scenario s $[kWh]$
- $E_{-}^{h,s} \in \mathbb{R}^+$: charge energy for the storage system at hour h in the scenario s $[kWh]$
- $LOC^{h,s} \in \mathbb{R}^+$: level of charge of the battery at hour h in the scenario s $[kWh]$
- $P_{sold}^{h,s} \in \mathbb{R}^+$: energy sold to the grid at hour h in the scenario s $[kWh]$
- $P_{shared}^{h,s} \in \mathbb{R}^+$: shared energy at hour h in the scenario s $[kWh]$
- $P_{+}^{n,h,s} \in \mathbb{R}^+$: energy surplus at hour h in the scenario s for building n $[kWh]$
- $P_{-}^{n,h,s} \in \mathbb{R}^+$: energy deficit at hour h in the scenario s for building n $[kWh]$
- $P_{+,EC}^{h,s} \in \mathbb{R}^+$: total energy surplus of the REC at hour h in the scenario s $[kWh]$
- $P_{-,EC}^{h,s} \in \mathbb{R}^+$: total energy deficit of the REC at hour h in the scenario s $[kWh]$

Auxiliary variables
These binary auxiliary variables are used in the optimization model to formulate big-M constraints:

- $is_{surplus}^{n,h,s} \in \{0,1\}$: binary variable of value 1 if there is a surplus at hour h in the scenario s for building n, 0 if there is a deficit
- $N_{charged}^{h,s} \in \mathbb{N}$: integer variable representing the number of fully charged batteries at hour h in the scenario s
- $is_{immission}^{h,s} \in \{0,1\}$: binary variable of value 1 if the shared energy - the minimum between immission and withdrawal - at hour h in the scenario s is equal to the immission $P_{sold}^{h,s}$, 0 if it is equal to the withdrawal $P_{grid}^{h,s}$

Cost and technical parameters

- $CAPEX_{PV}$: total acquisition and installation cost for a photovoltaic unit [€/kW]
- $CAPEX_S$: total acquisition and installation cost for a storage unit [€/kWh]
- $OPEX_{PV}$: operational cost for a photovoltaic unit in one year [€/kW]
- $OPEX_S$: operational cost for a storage unit in one year [€/kWh]
- c_+ : price for energy withdrawn from the grid [€/kWh]
- c_-^h : price for the sale of energy to the grid at hour h [€/kWh]
- p_{shared} : incentive for shared energy [€/kWh]
- r : interest rate [%]
- $\hat{P}_{PV}$: maximum PV energy product by a PV unit [kWh]
- C_{max}: maximum battery capacity [kWh]
- x_{PV}^{max}: maximum number of installable PV units
- x_S^{max}: maximum number of installable batteries

Uncertain parameters

- $L^{n,h,s}$: energy load at hour h in the scenario s for building n [kWh]
- $P_{PV}^{n,h,s}$: energy produced by a PV unit at hour h in the scenario s for building n [kWh]

Objective Function

The objective function of the problem is chosen to be the Net Present Value (NPV) of the system, defined as the difference between the present value of cash inflows and the present value of cash outflows over a period of time. In this case, we will consider a period of time of 20 years. This metric is similar to the widely used annualized total cost, which, instead, computes the cost per year of owning and operating an asset over its entire lifespan, but it is more adopted for practical purposes. An important index closely related to the NPV is the payback time, i.e., the is the amount of time that it takes for the initial cost of a project to equal to the discounted value of expected cash flows. The formula of the NPV includes the

first stage investment cost and the second stage operational costs. We suppose that lifetime of the photovoltaic system to be 20 years, while for the storage system to be 10 years. This implies that, after 10 year, it will be necessary to substitute the batteries. Such cost has to be actualized, i.e., discounted to the year zero. Therefore, the total investment cost is computed as

$$INV = CAPEX_{PV} \sum_{n=1}^{N} x_{PV}^{n} + CAPEX_S x_S + SUB_S x_S \frac{1}{(1+r)^{10}}, \tag{4.2}$$

where r is the (constant) rate of interest. The total operational cost in a year, instead, depend on the scenario and is computed as the sum of all the costs and revenues due to the operational variables in a year, by the following formula:

$$OP_s = MAIN + K\Big(- \sum_{h=1}^{H} c_{-} P_{grid}^{h,s} + c_{+}^{h} P_{sold}^{h,s} + p_{shared} P_{shared}^{h,s} \Big), \tag{4.3}$$

where $K = 30.417$, is a constant necessary to pass from the twelve days considered in each scenario to a year, and $MAIN$ is the total annual maintenance cost of the system, given by

$$MAIN = OPEX_{PV} \sum_{n=1}^{N} x_{PV}^{n} + OPEX_S x_S \tag{4.4}$$

Summing up all the terms, and considering the lifetime of 20 years, the NPV of the investment at 20 years, will be given by

$$NPV = INV + \sum_{t=0}^{20} \frac{1}{(1+r)^t} \bar{OP}, \tag{4.5}$$

where

$$\bar{OP} = \frac{1}{S} \sum_{s=1}^{S} OP_s,$$

is the average total annual operational cost, based on the different scenarios.

4.2.2 Formulation and constraints

Optimization model

$$\min \quad NPV \tag{4.6a}$$

$$\text{s.t.} \quad x_{PV}^{n} \leq \hat{x}_{PV}^{n}, \quad n \in \mathcal{N} \tag{4.6b}$$

$$x_S \leq \hat{x}_S \tag{4.6c}$$

$$LOC^{h,s} \leq x_S C_{max}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6d}$$

$$LOC^{h,s} \geq 0.1 x_S C_{max}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6e}$$

$$\frac{1}{-}E_{-}^{h,s} \leq 0.5 x_S C_{max}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6f}$$

$$\frac{1}{\eta}E_{-}^{h,s} \leq LOC^{h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6g}$$

$$\eta E_{+}^{h,s} \leq 0.5 x_S C_{max} \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6h}$$

$$\eta E_{+}^{h,s} \leq x_S C_{max} - N_{charged} C_{max}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6i}$$

$$P_{+}^{n,h,s} - P_{-}^{n,h,s} = x_{PV}^{n} P_{PV}^{n,h,s} - L^{n,h,s}, \quad n \in \mathcal{N},\ h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6j}$$

$$P_{+,EC}^{h,s} = \sum_{n=1}^{N} P_{+}^{n,h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6k}$$

$$P_{-,EC}^{h,s} = \sum_{n=1}^{N} P_{-}^{n,h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6l}$$

$$P_{+,EC}^{h,s} = E_{+}^{h,s} + P_{sold}^{h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6m}$$

$$P_{-,EC}^{h,s} = E_{-}^{h,s} + P_{grid}^{h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6n}$$

$$LOC^{0,s} = 0.1 x_S C_{max}, \quad s \in \mathcal{S} \tag{4.6o}$$

$$LOC^{h,s} = LOC^{h-1,s} - \frac{1}{\eta}E_{-}^{h-1,s} + \eta E_{+}^{h-1,s}, \quad h \in \mathcal{H} - \{0\},\ s \in \mathcal{S} \tag{4.6p}$$

$$P_{+}^{n,h,s} \leq is_{surplus}^{n,h,s} M_1, \quad n \in \mathcal{N},\ h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6q}$$

$$P_{-}^{n,h,s} \leq (1 - is_{surplus}^{n,h,s}) M_1, \quad n \in \mathcal{N},\ h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6r}$$

$$P_{shared}^{h,s} \leq P_{grid}^{h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6s}$$

$$P_{shared}^{h,s} \leq P_{sold}^{h,s}, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6t}$$

$$P_{shared}^{h,s} \geq P_{grid}^{h,s} - is_{immission}^{h,s} M_2, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6u}$$

$$P_{shared}^{h,s} \geq P_{sold}^{h,s} - (1 - is_{immission}^{h,s}) M_2, \quad h \in \mathcal{H},\ s \in \mathcal{S} \tag{4.6v}$$

The first two constraint of the optimization model are maximum capacity constraints for the PV and storage size. Constraints (4.6d) imposes that the level of charge of the battery at each hour cannot exceed the maximum installed capacity, while constraint (4.6e) imposes that the level of charge of the battery can go below a threshold, set to the 10% of the maximum capacity. Constraints (4.6f) and (4.6g) regulate the discharge of the battery: each hour it cannot be exceed the present level of charge and the half of the maximum capacity. Similarly, constraint (4.6h) states that the charge cannot be more the half of the maximum capacity, while (4.6i) imposes that it is not possible to charge more than the capacity of the not-fully-charged batteries. Constraint (4.6j) balances the energy flows at a prosumer level: for each hour and for each building, if the generation is higher

than the energy demand, the difference is an energetic surplus, otherwise a deficit. These surplus and deficit variables will be considered at a community level, as accounted by constraints (4.6k) and (4.6l). Constraint (4.6m) is an operational constraint that manage the energy surplus of the community: at each hour, it can either spent to charge the battery or it can be sold to grid. Similarly constraint (4.6n) concerns the deficit of the community, which can covered through discharging or withdrawing from the grid. Constraints (4.6o) and (4.6p) involve the dynamic update of the level of charge: the first set the initial level of charge of the battery to the 10% of the total capacity, while the second equation states that the new level of charge is given by adding the charged energy to the level of charge of the previous hour and subtracting the discharged energy. Constraints (4.6q) and (4.6r) represent the big-M constraint for the mutual exclusivity of $P_{+}^{n,h,s}$ and $P_{-}^{n,h,s}$, i.e., at each hour and for each building, either $P_{+}^{n,h,s}$ is non-negative and $P_{-}^{n,h,s}$ is zero or vice versa. The last four constraints, instead, express that $P_{shared}^{h,s}$ is equal to the minimum between $P_{sold}^{h,s}$ and $P_{grid}^{h,s}$.

Chapter 5

Model Evaluation

In this section, we will evaluate the model before presenting the results. At first we will analyze the convergence of the model to the optimal solution, in terms of number of scenarios necessary to reach predetermined levels of confidence and error on the mean. Secondly, we will discuss about which parameters are more influential to the final solution and the design provided by the stochastic problem differs depending on the uncertainties considered in the model.

5.1 Convergence Results

Being the solution based on a Monte Carlo simulation, it is important to understand which is the sufficient amount of Monte Carlo samples that leads to a satisfactory approximation of the real solution, considering to find a reasonable trade-off between accuracy and computational times. Bashir and Sadeh proposed in [21] some Monte Carlo simulation stopping rules, i.e., ways to determine how many simulations are necessary to make sure that the mean value estimator lies within a certain tolerance from the real mean. Consider the output of the simulation as a random variable $X \in \mathbb{R}$ and define $X_i, i \in \mathbb{N}$ as i.i.d. copies of it. As X is defined as the output of a simulation we can assume that $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$ are finite real numbers. consider the empiric mean

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} X_i, \tag{5.1}$$

and the corresponding error $|\bar{X}_N - \mu|$, which itself is a random variable. Therefore, we cannot ask for a useful upper and lower bound for it, but seek for a confidence interval with predefined width δ and confidence level α so that

$$\mathbb{P}(|\bar{X}_N - \mu| \leq \delta) \geq \alpha. \tag{5.2}$$

The aim is to find a formula that describes the relation between sample size N, probability α and allowed absolute error δ, we are able stop the Monte Carlo simulation whenever the allowed error δ and the failure probability α are sufficiently small. The most frequently used stopping rule is based on the Central Limit Theorem (CLT)

$$\frac{\bar{X}_N - \mu}{\sqrt{N}\sigma} \xrightarrow[N\to\infty]{P} Y \sim \mathcal{N}(0,1). \tag{5.3}$$

From the CLT we obtain the following asymptotic Gauss-Distribution Stopping Rule

$$N \geq \left(\frac{z_{\alpha/2}\sigma}{\delta}\right)^2, \tag{5.4}$$

where $z_{\alpha/2}$ is the quantile of level $\alpha/2$ of the standard Gaussian distribution.

To determine the necessary number of simulations, we generated 1000 scenarios, considering all the uncertainties described in the previous sections, and applied the stopping criterion to the obtained values of the objective function. Table 5.1 shows the computed number of Monte Carlo simulations for some levels of confidence α and for different errors from the mean δ.

	10%	5%	1%
0.95	9	37	233
0.975	12	49	304
0.99	16	64	402

Table 5.1: Number of simulations ensuring different levels of confidence (rows) and levels of error from the mean (columns).

Depending on the desired levels of confidence and error, one can decide the number of simulations to use for running the model. Clearly, the more simulations will be performed, the highest the computational time will be. Figure 5.1 shows the total computational time needed to solve the Sample Average Approximation problem with a different number of simulations. Being the the problem mixed-integer, we consequently notice an exponential relation between the number of simulations and the computational time. In choosing the number of simulations, is, thus, necessary to find a trade-off between the accuracy of the solution and the computational effort.

5.2 Value of Stochastic Information

In this section we aim to understand if the introduction of the stochasticity into the model leads to improvements to the final solution, enough to justify the high

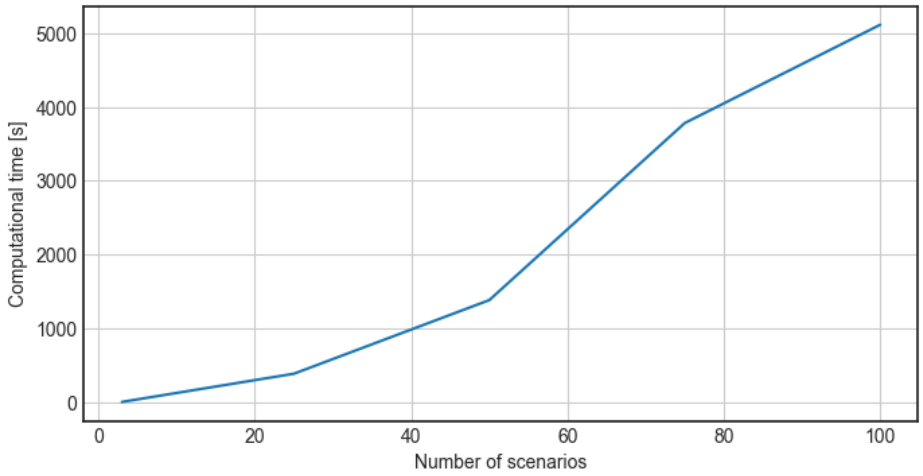

Figure 5.1: Computational time for different numbers of scenarios.

raise in complexity of the problem. And, how much the uncertainty in a parameter affect the design of the REC. To answer these questions, we propose a qualitative analysis aiming to evaluate how the solution of the deterministic model perform, considering various scenarios of the uncertain parameters.

To begin with, we consider the influence of the PV generation parameters on the final solution. To do so, based on the historical data, we generated three scenarios of PV production, namely high, average and low, as shown in Figure 5.2, and run the deterministic separately for each of these scenarios, fixing the other parameters to their mean values.

Figure 5.3 shows the results of these runs. We notice that the solution varies a lot between these scenarios: the case of high PV generation leads to huge difference in the total kW of PV installed, producing an increase in the NPV and self-sufficiency, while lowering the level of self-consumption of REC. This suggests that the introduction of the uncertainty in the generation is essential and can strongly influence the final output of the model. In Figure 5.4, instead, we evaluated how the selected KPIs varies when solving the deterministic model on randomly generated PV generation profiles as described in Chapter 4. The deterministic solution, depicted with the orange dot, seems to be consistent with the most common value of PV sizing, despite some extreme cases. However, it differs in terms of the others economic and energetic KPIs, possibly due to the non-smooth randomly generated PV profiles compared to the typical average profiles.

As to the impact of modelling the uncertainty in the energy load, we considered

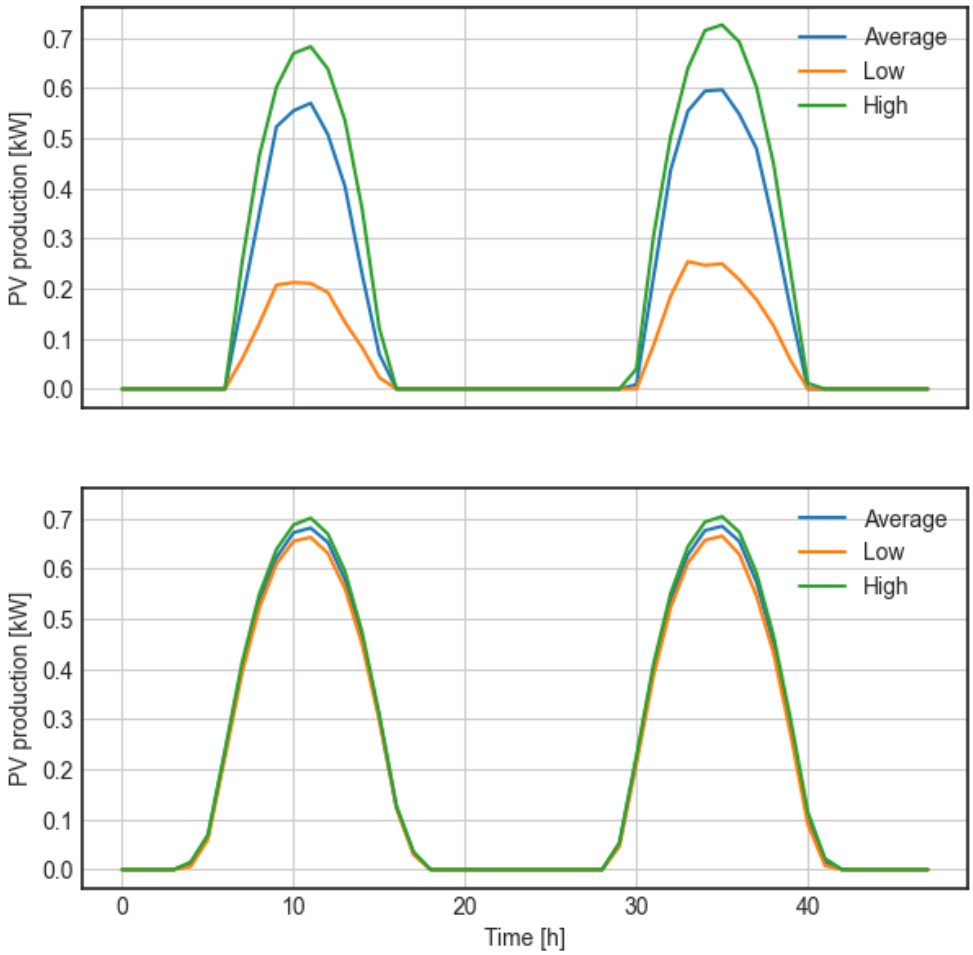

Figure 5.2: Sizing and KPIs for the three test scenarios of the magnitude of the PV generation.

two aspect to be evaluated. At first, we analyzed the magnitude of the energy demand, considering three scenarios starting from the simulated average load, and scaling with two constant factors, as shown in Figure 5.5. Secondly, we tested the uncertainty modelling of the composition of the residential buildings. In particular, we considered three scenarios in which we suppose that each building is composed by only units of a particular social group. We tested the solution examining the social groups YB, LI and LF. The energy load of the class YB exhibit a single-peak profile, while the others are characterized by a double-peak profile, with different magnitudes.

Contrary to the PV generation, in the case of the uncertainty in the energy

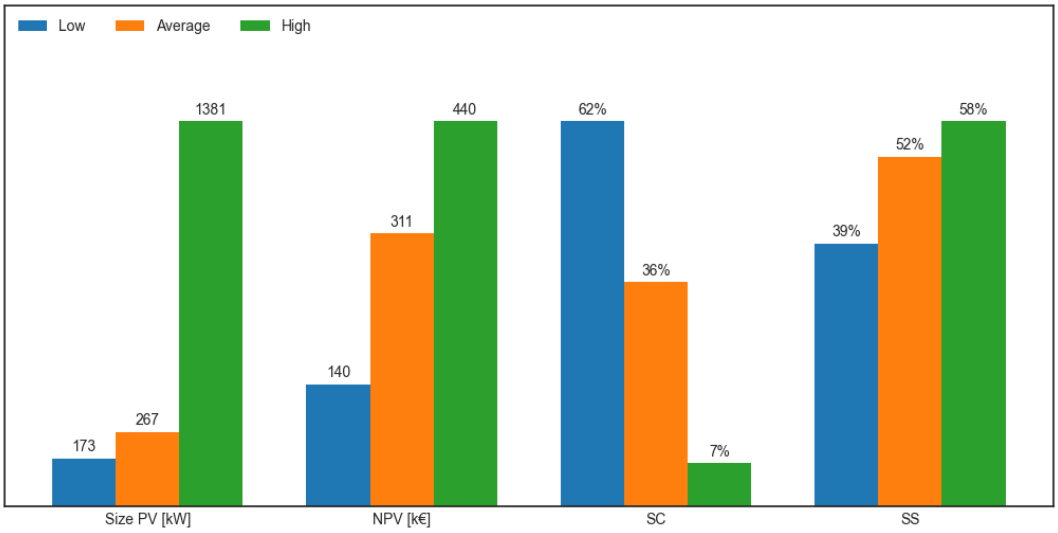

Figure 5.3: Sizing and KPIs for the three test scenarios of the magnitude of the PV generation.

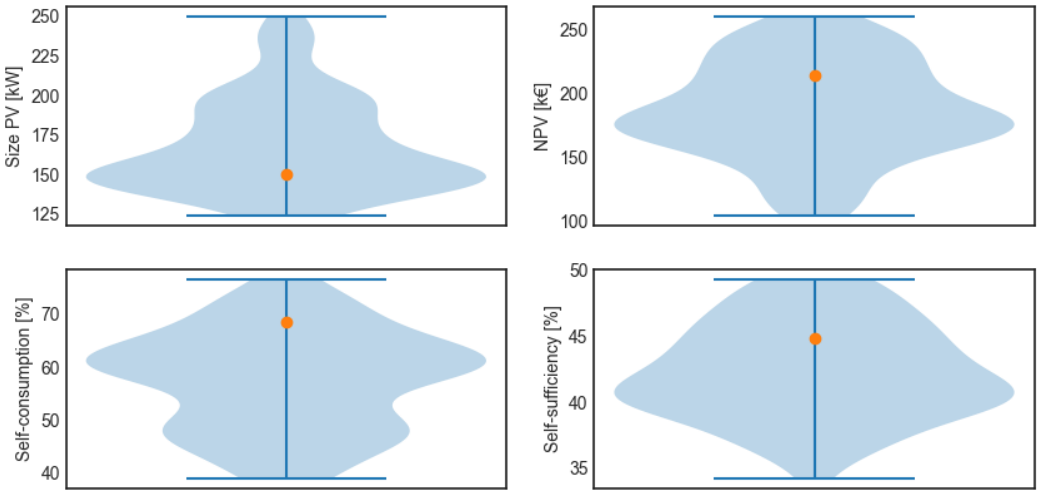

Figure 5.4: Violin plots of the sizing and KPIs for the three test scenarios of the magnitude of the PV generation.

load the differences between the three scenarios are less harsh. These results are displayed in Figure 5.6. Increasing the total energy load leads to an increase of the number of photovoltaic units installed quite linearly. On the simulated scenarios in Figure 5.7, we assist to a difference up to a maximum of the 70% in the PV sizing, of the 50% in the NPV, of the 35% in the self-sufficiency and of the 30% in the self-consumption, compared to the deterministic solution. As to the uncertainty

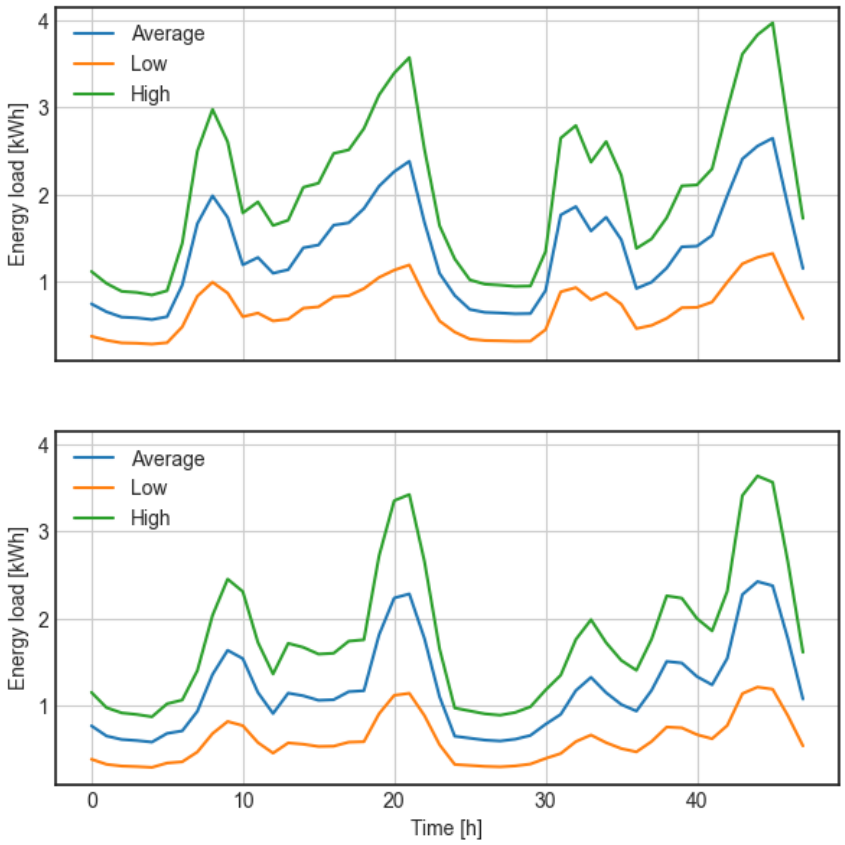

Figure 5.5: Sizing and KPIs for the three test scenarios of the magnitude of the PV generation.

in the composition in the residential buildings. In particular in Figure 5.9 we run the deterministic model considering different scenarios of configurations of the residential buildings.

Finally, we evaluated the impact of the energy prices on the output of the optimization model. The three scenarios have been constructed used the values of the PUN and bill described in Chapter 4. Being the objective function of the problem an economic metric, as expected, energy princes have a strong influence on the solution. As matter of fact, as shown in Figure 5.10 passing from the scenario of average prices to the scenario of low prices produces a variation of the PV sizing of 57 kW, resulting in a considerable variation of the NPV. On the simulated scenarios in Figure 5.11, we assist to a difference up to a maximum of the 75% in the PV sizing, of the 42% in the NPV, of the 38% in the self-sufficiency and of the 31% in the self-consumption, compared to the deterministic solution.

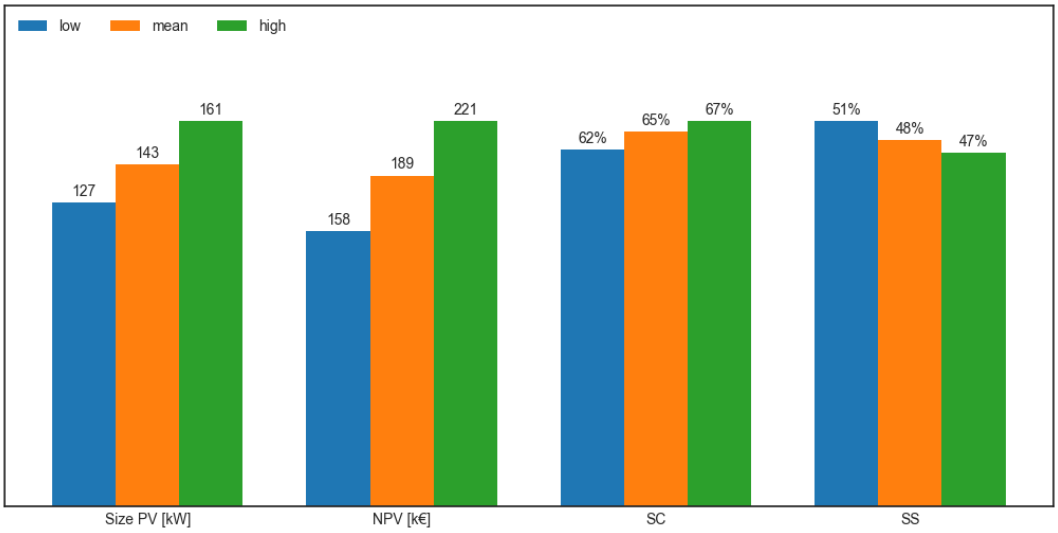

Figure 5.6: Sizing and KPIs for the three test scenarios of the magnitude of the energy load.

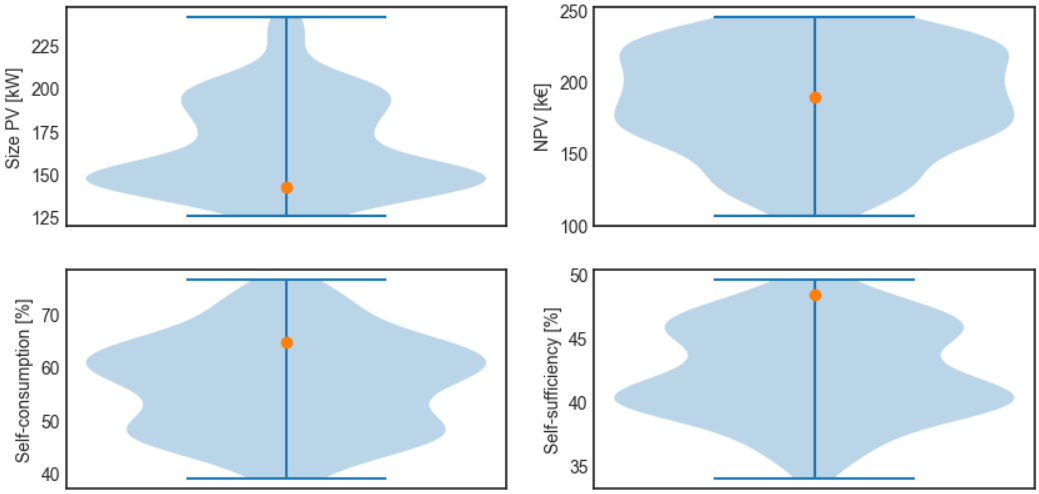

Figure 5.7: Violin plot of the sizing and KPIs for the three test scenarios of the magnitude of the energy load.

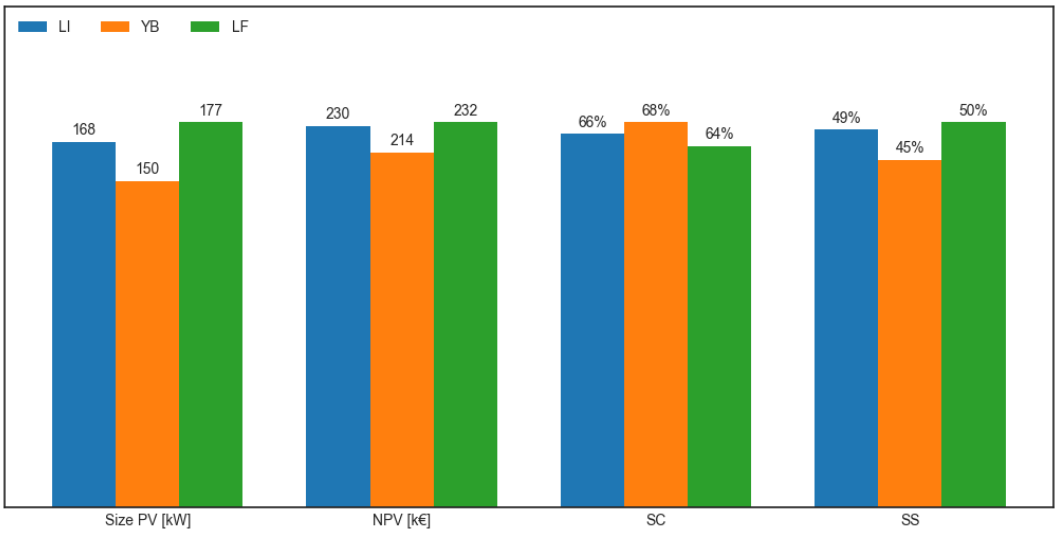

Figure 5.8: Sizing and KPIs for the three test scenarios of the residential building composition.

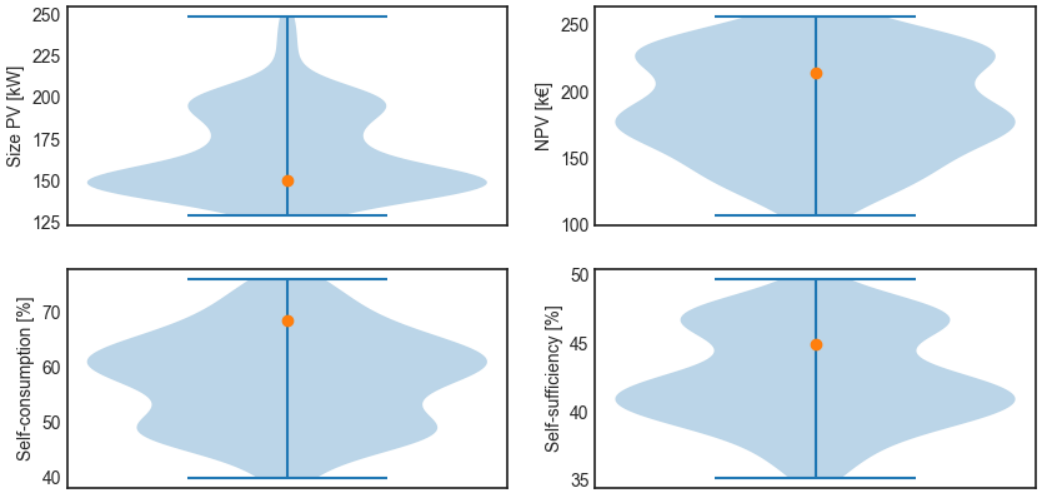

Figure 5.9: Violin plot of the sizing and KPIs for the three test scenarios of the residential building composition.

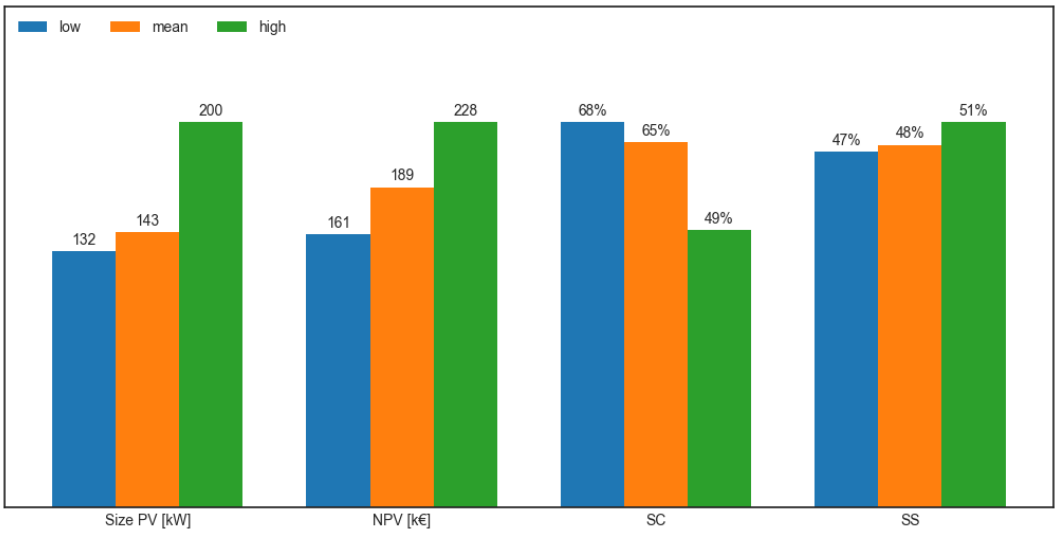

Figure 5.10: Sizing and KPIs for the three test scenarios of the energy prices.

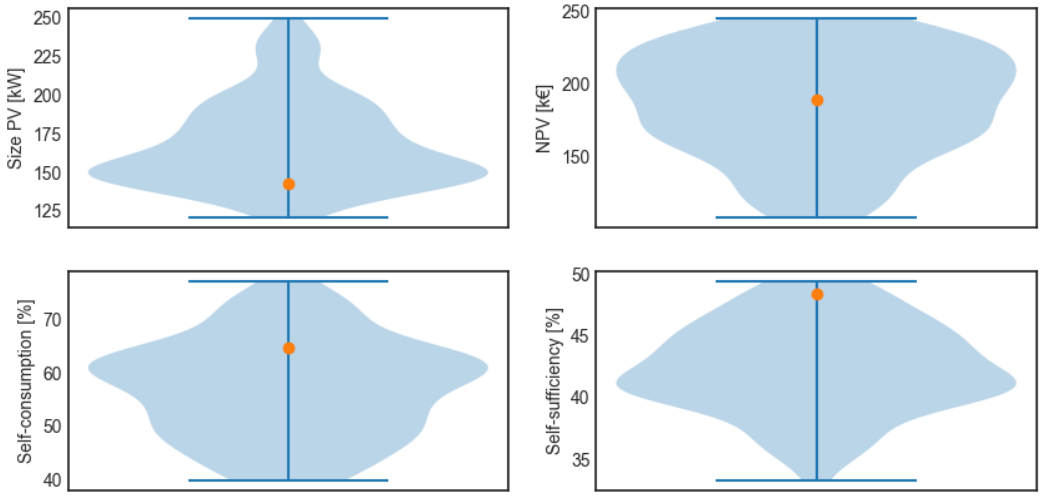

Figure 5.11: Violin plot of the sizing and KPIs for the three test scenarios of the energy prices.

Chapter 6

Analysis of Results

This chapter presents the results of the optimization model resolution for various REC configurations. The first section analyses the variation of sizing and other economic and energetic KPIs for different configurations. The second section discusses the comparison between the stochastic solution and the deterministic one, which is obtained by solving the problem considering the average values of the uncertain parameters.

6.1 Sizing and KPIs

To assess the performance of the optimization model, we will evaluate the deterministic and stochastic model's results in terms of optimal sizing and economic and energetic KPIs. These include the NPV (the objective function of the problem), payback time, self-sufficiency, and self-consumption. Additionally, we will determine whether the installation of the storage system is cost-effective, given the pre-defined investment and operational prices of the batteries. As mentioned in Chapter 4, the area of analysis is the city of Taranto. Therefore, the PV generation and energy demand parameters refer to this case study. Three different configurations will be considered, varying the composition of community members and cost parameters. The results will be compared between them and against the deterministic case. These three configurations are chosen as follows:

- **Configuration 1**: In this configuration we consider 10 residential buildings, 5 offices, 3 restaurants, 2 schools and 1 store, where we suppose that 8 residential buildings are only consumers, thus, they cannot install PV panels. As prices for the storage systems, we set $CAPEX_S = 1400$ €$/kWh$ and $OPEX_S = 80$ €$/kWh$.
- **Configuration 2**: In this configuration we increase the number of residential

buildings in the REC, going from 10 to 20, and supposing that only 4 of them are prosumers, while keeping the other parameters equal to configuration 1. This is done specifically to evaluate how the shared energy varies in a configuration as the number of PODs increases and how it turns to self-sufficiency and self-consumption.

- **Configuration 3**: In this configuration we will keep the composition of the REC as in configuration 1, but we vary the installation and operational costs of the storage systems, in order to present a situation in which the storage system is actually installed. We set $CAPEX_S = 400$ €$/kWh$ and $OPEX_S = 8$ €$/kWh$.

The ohter parameters will remain the same for all three configurations. In each case, profiles for PV generation and energy load, as well as scenarios for energy prices, will be randomly generated as described in Chapter 4. Building parameters, such as azimuth and slope, will be initialized randomly. For the stochastic simulations, we set the number of Monte Carlo samples to 100, which ensures a confidence level of 99% with an error from the mean value of no more than 5%. Besides that, we will set the following parameters:

- $CAPEX_{PV} = 1600$ €$/kW$,
- $OPEX_{PV} = 20$ €$/kW$,
- $\hat{P}_{PV} = 1\ kW$,
- $p_{shared} = 0.11$ €$/kWh$,
- $r = 0.03\%$.

We will start by analyzing the results for Configuration 1. Table 6.1 shows the outcomes of the deterministic and stochastic problems considering the parameters chosen for this configuration. In the third row of the table, the relative differences of the KPIs of the stochastic solution with respect to the deterministic one are computed. It can be immediately noticed that, in both cases, the battery is not installed, due to its high costs and since it is more convenient to import the surplus energy of the community into the grid, instead of charging the storage system. In the stochastic solution the sizing of the PV system varies by approximately 12% compared to the deterministic case. This variation can be justified by the results obtained in Chapter 5. As matter of fact, under the scenario of high PV generation there is a significant increase in the total PV installation compared to the average scenario. On the other hand, a lower variation is observed when moving from the average scenario to the low one, as shown in Figure 5.3.In addition, the introduction

of uncertainty in energy prices in the stochastic model has a significant impact on the final output, being the objective function of the problem economic in nature. The evaluated economic and energetic KPIs remain similar for both the deterministic and stochastic models. Regarding economic performance, the stochastic solution results in an increase in initial investment cost and NPV of 7% compared to the deterministic model. In both cases, the investment pays back in 10 years. Figure 6.1 shows the NPV over a 20-year planning horizon for the stochastic solution. The orange bar in the figure represents the initial investment of € 256,000, which is recouped after 10 years. In terms of energetic KPIs, the deterministic solution outperforms the stochastic one, resulting in an approximately 8% reduction in self-sufficiency, 2% reduction in self-consumption, and 2% reduction in shared energy percentage. One possible reason for this is that in the deterministic case, the optimization only considers the mean profiles of PV production and energy load. As a result, fluctuations in these profiles and their interactions are not taken into account.

KPI	Deterministic	Stochastic	Difference
Size PV [kW]	143	160	+12%
Battery capacity [kWh]	0	0	
NPV [€]	177186	189151	+7%
Investment [€]	238400	256000	+7%
Payback time	10	10	
Self-sufficiency	52.17%	44.53%	-8%
Self-consumption	66.87%	64.46%	-2%
Shared energy	55.43%	53.88%	-2%

Table 6.1: Optimal sizing, economic and energetic KPIs of the deterministic and stochastic solutions for Configuration 1.

Here, we examine the solution obtained from the stochastic model for Configuration 1. Figure 6.2 displays the unitary PV generation profiles for the various combinations of azimuth and slope of the buildings that make up the REC. It is evident that all the PV capacity is installed on a single building, namely on the school with azimuth angle $\alpha = 127°$ and slope angle 39°. The motivation behind this is that with these parameters we achieve a high level in magnitude of PV production. towards the afternoon hours. This better matches the energy demand of REC, as shown in Figure 6.3, leading to a greater percentage of shared energy and an increase in NPV under the given incentive scheme.

Figures 6.4 and 6.5 display the optimal energy dispatch operations for January and February, and for July and August, respectively. In January and February, the

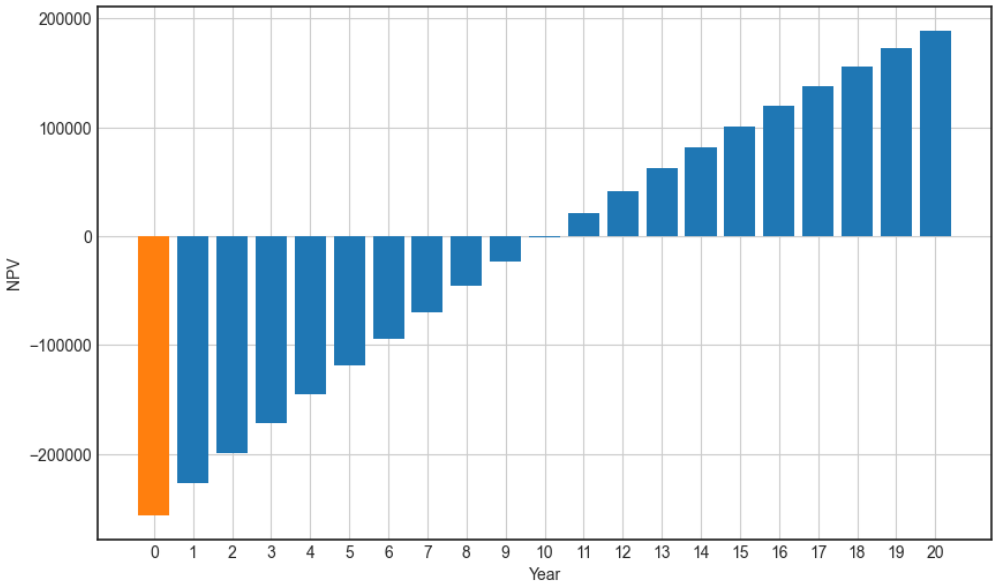

Figure 6.1: NPV over a time horizon of 20 years of the stochastic solution for Configuration 1.

total PV production is lower, resulting in a higher amount of energy withdrawal from the grid and an increase in operational costs. Additionally, the percentage of shared energy compared to energy imported into the grid is higher. During July and August, there is no physical self-consumption of energy in the school where the PV is installed since there is no energy demand for building where all the PV capacity is installed. As a result, all the PV production is either used as shared energy or imported into the grid. Additionally, during the summer, it is almost entirely possible to cover the energy demand of the REC through shared energy.

Table 6.2 compares the results of the stochastic solutions for the three configurations described above. Configuration 2 differs from Configuration 1 in the number of residential buildings in the community, which is doubled. As a result, the total installed PV increases by 30 kW. From an energy perspective, the increase in the number of REC consumers leads to a rise in total energy demand. As a result, Configuration 2 experiences a decrease in self-sufficiency and an increase in self-consumption and shared energy percentage. The change in community size makes it impossible to compare the NPV of the first two configurations. Nevertheless, both configurations have a payback period of 10 years. In Configuration 3, installation, substitution, and maintenance costs for the storage system are lower than those in Configuration 1. As a result, we install 40 kWh of storage capacity and simultaneously increase the PV capacity by 5 kW to better utilize the battery discharge to

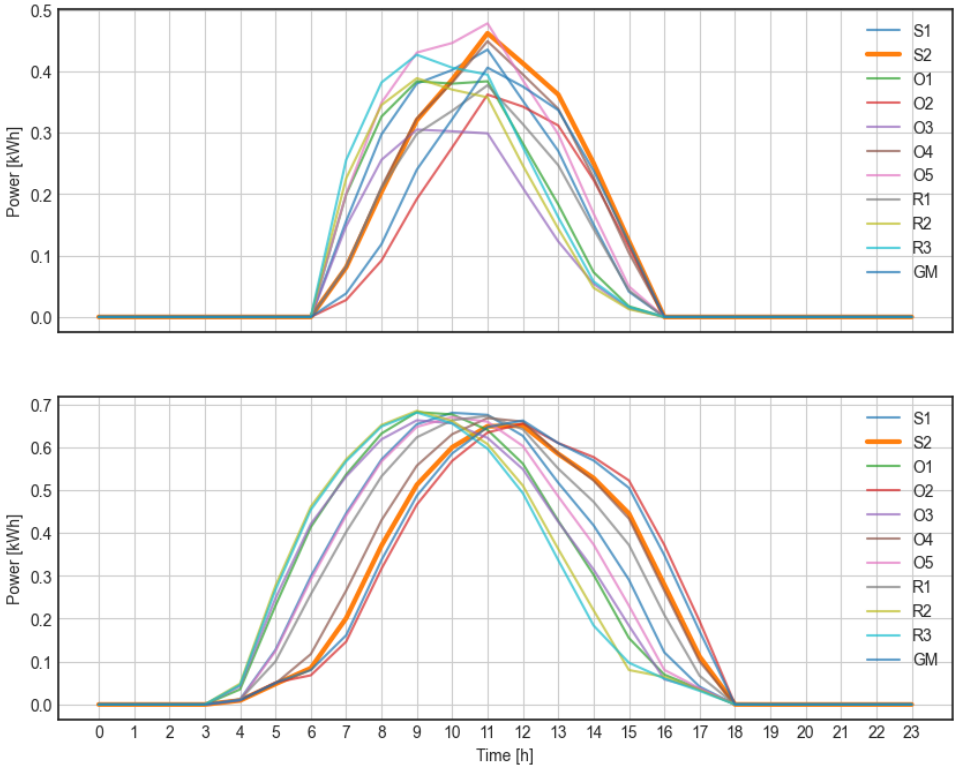

Figure 6.2: Average unitary PV generation of the different buildings, in January (top) and July (bottom).

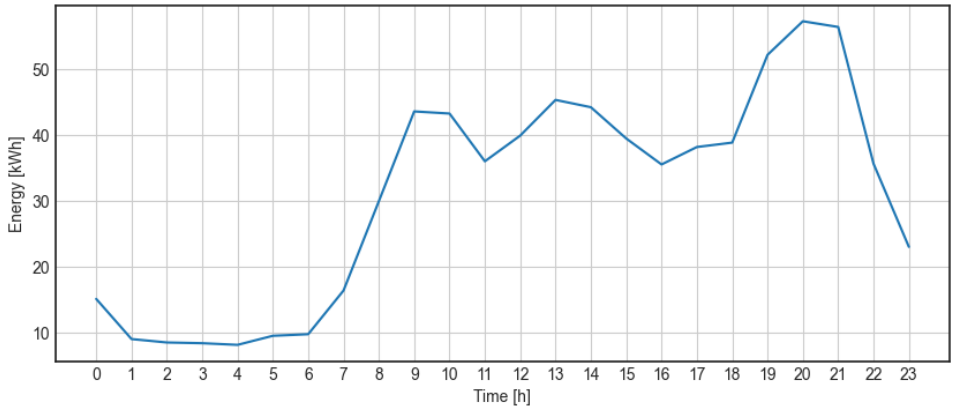

Figure 6.3: Average total energy load of the REC in January.

meet the energy deficits of the REC. The installation of the storage system leads to a significant increase in the initial investment. However, after 20 years, the

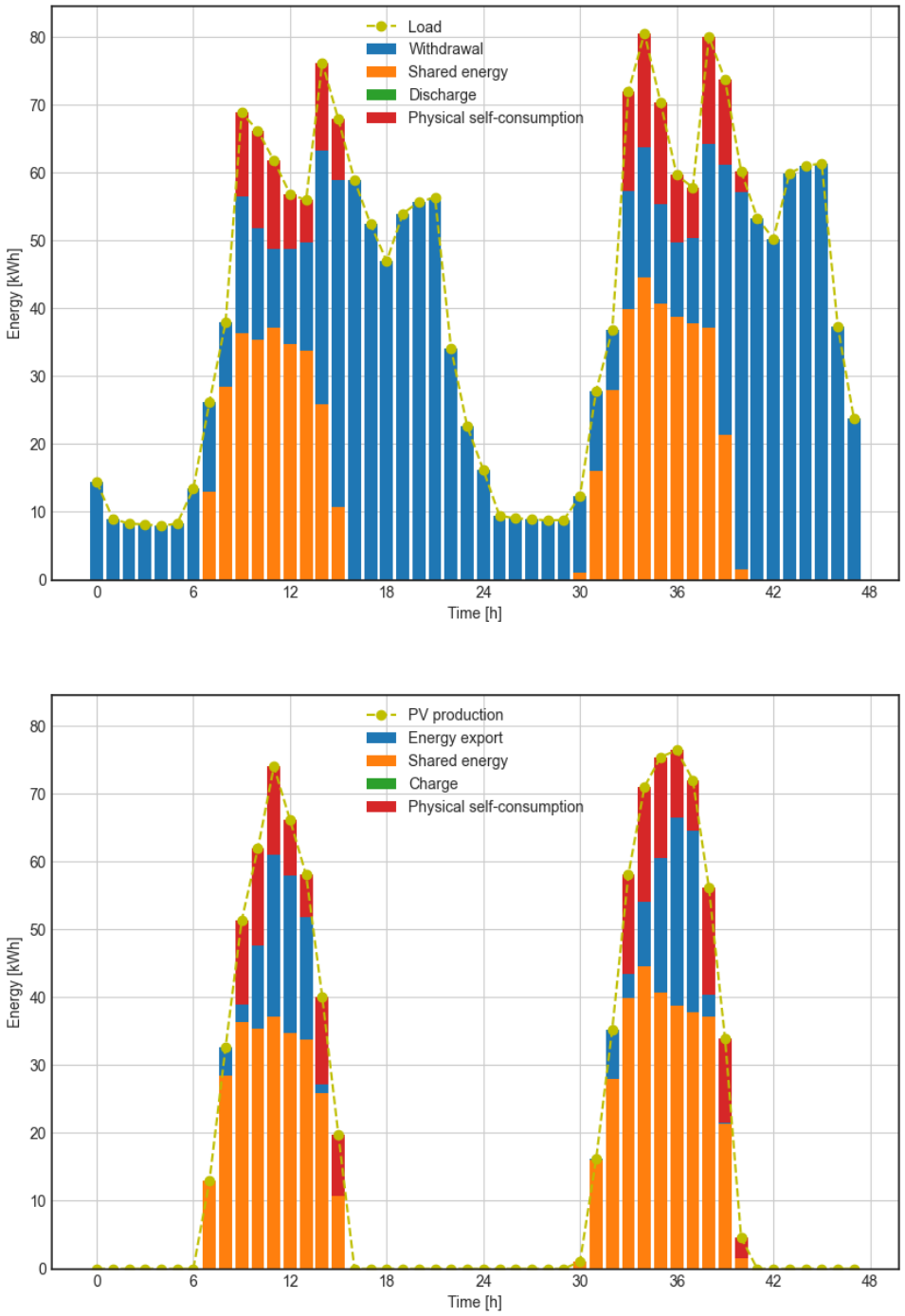

Figure 6.4: Average optimal energy dispatch operations for the stochastic solution of Configuration 1 in January-February.

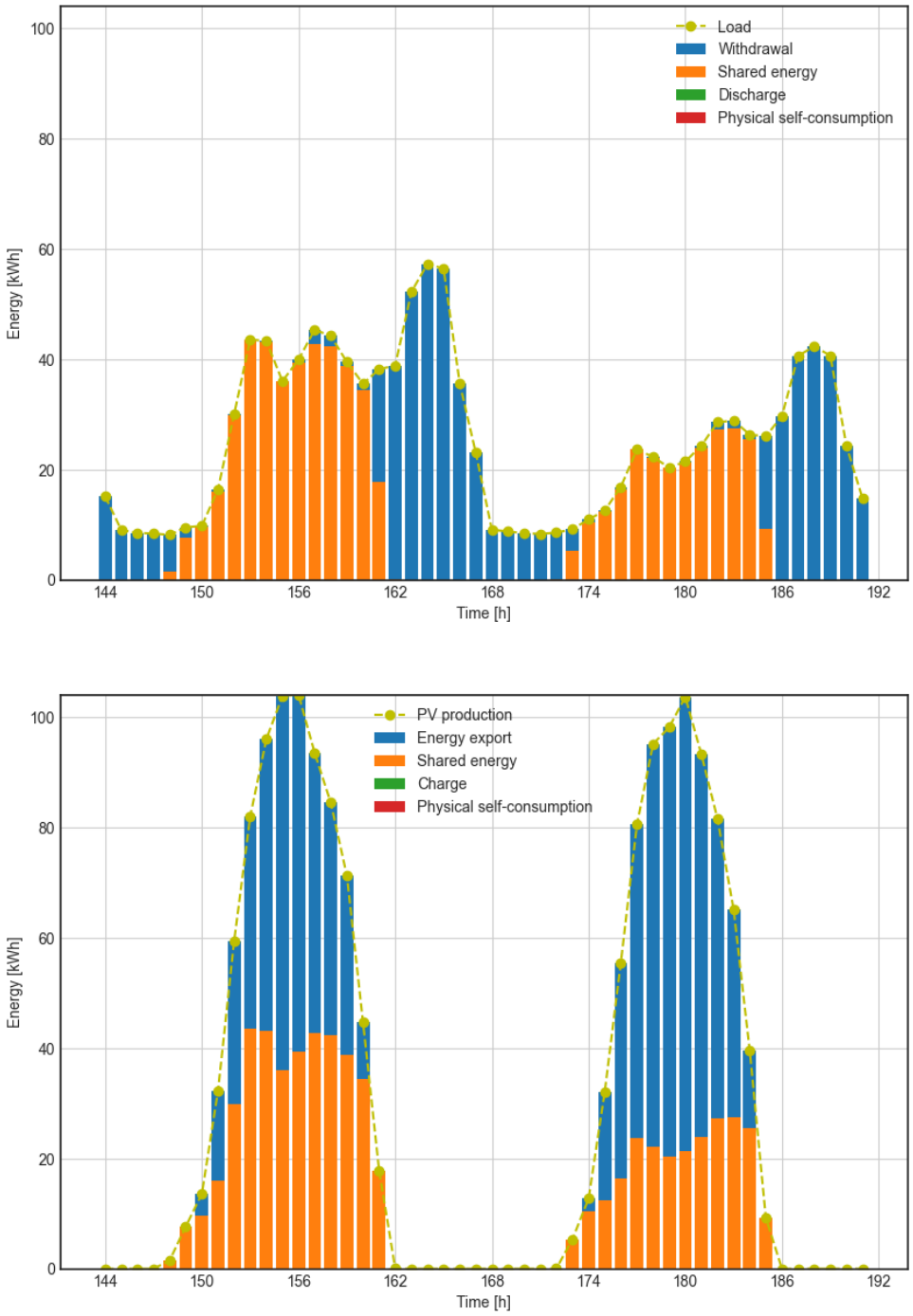

Figure 6.5: Average optimal energy dispatch operations for the stochastic solution of Configuration 1 in July-August.

net present value (NPV) of both configurations is almost identical. This indicates that the configuration with the storage system has lower operational costs and is therefore more convenient in the long term. Additionally, both the self-sufficiency and self-consumption of Configuration 3 increase compared to Configuration 1, as the charged and discharged energy are factored into the computation of these metrics. The REC is able to achieve approximately a promising 55% self-sufficiency and over 80% self-consumption with the storage system. Finally, in Figure 6.6 and 6.6 show the optimal energy dispatch operations on average for January and February and for July and August respectively in the case of Configuration 3.

	1	2	3
KPI	Configuration		
Size PV [kW]	160	190	165
Battery capacity [kWh]	0	0	40
NPV [€]	189151	230121	203001
Investment [€]	256000	304000	2859523
Payback time	10	10	11
Self-sufficiency	44.53%	37.81%	55.16%
Self-consumption	64.46%	71.82%	81.79%
Shared energy	53.88%	63.87%	43.53%

Table 6.2: Comparison of the optimal sizing, economic and energetic KPIs of the stochastic solutions for the three configurations.

6.2 Analysis of the Performance

In this section, we will test and compare the performance of the stochastic solution with the performance of the deterministic solution. To do so, we will run the stochastic model fixing the first-stage variables, i.e., the sizings of the PV and storage systems, obtained by running the stochastic and the deterministic models. The second-stage variables will, instead, be left free. In particular, we will carry out this test by considering the deterministic and stochastic sizing obtained for Configuration 1 and reported in Table 4.6b. We will generate 1000 scenarios and evaluate how both the solutions behave with these new profiles of PV generation, energy load and energy prices. The aim is to understand how the solutions of the two models behave given perturbations in the initial parameters, thus, and to depict which of them is more stable.

Before analyzing the results of the described test, we verified if the the choice of

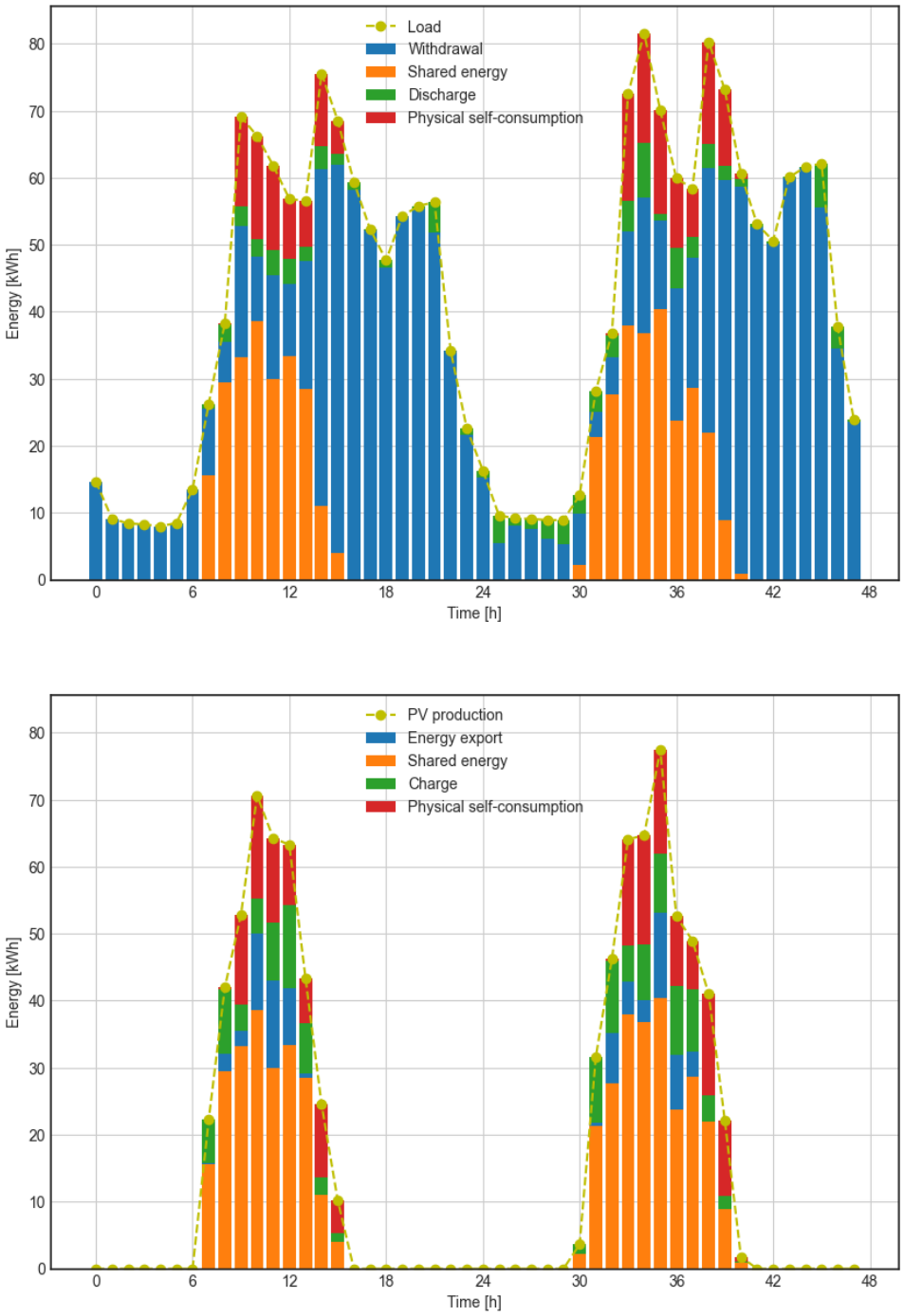

Figure 6.6: Average optimal energy dispatch operations for the stochastic solution of Configuration 3 in January-February.

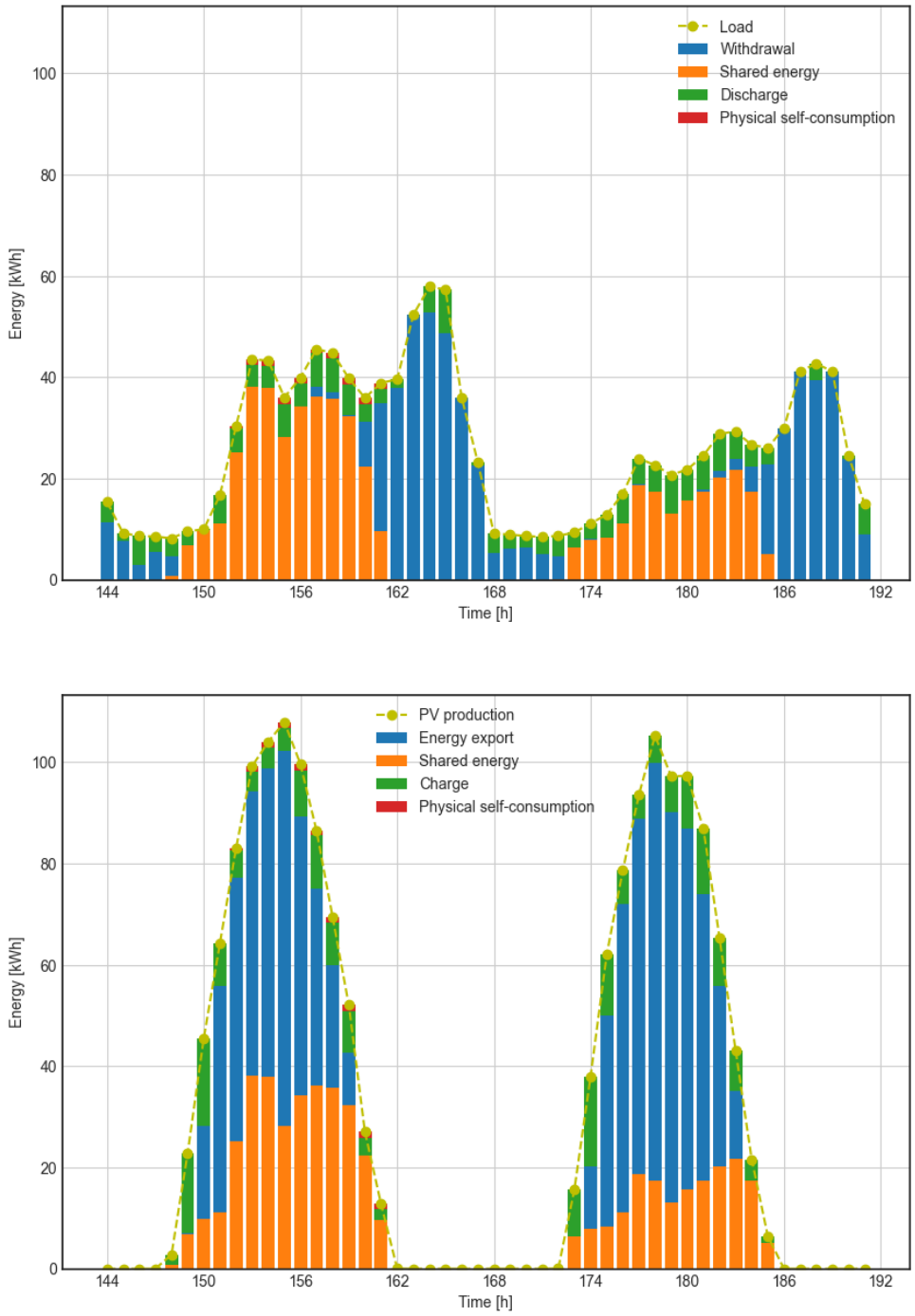

Figure 6.7: Average optimal energy dispatch operations for the stochastic solution of Configuration 3 in July-August.

Monte Carlo samples as described in Chapter 5 is actually producing a solution within a certain error from the mean value. In particular, for the simulations, a number of 100 scenarios has been chosen. Therefore, we expect to achieve, with a confidence level of the 99%, a maximum error of 5%, as described in Table 5.1. We run the stochastic model 50 times and sketched one boxplot for each of the KPIs under analysis, as shown in Figure 6.8, together with the values of the KPIs of the deterministic solution. In the top left plot, the results for the NPV are displayed. The NPV is the objective function of the optimization problem and the KPI used to determine the number of Monte Carlo simulations. In the plot the colored band is a visualization of the values of the NPV that differs from the mean NPV from the 50 simulations of the ±5%. The values of the NPV of the 50 simulations all fall inside the band, which means that the chosen number of Monte Carlo samples actually satisfies our expectations. Another important aspect to notice it that the value of the NPV of the deterministic solution is outside the ±5% error band. This implies that there is a statistically significant difference between the solution of the stochastic model and the solution of the deterministic model.

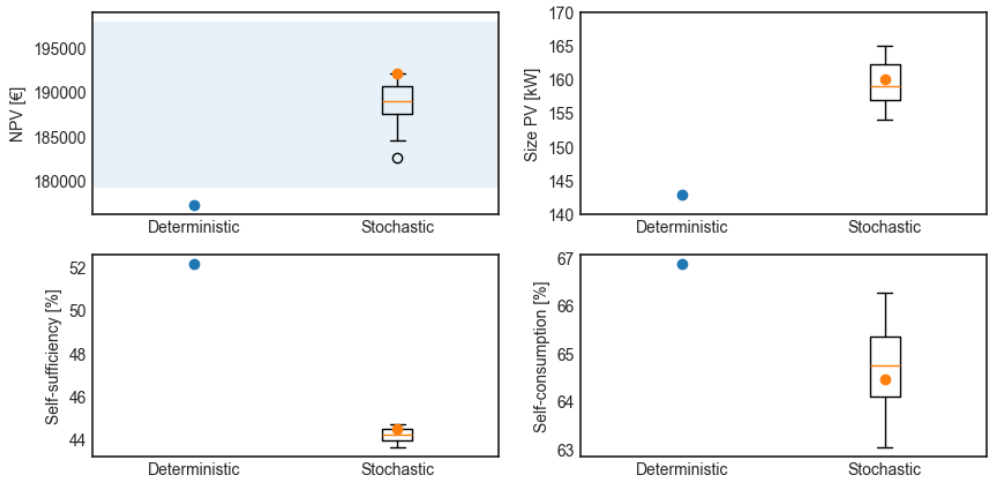

Figure 6.8: Comparison of the deterministic solution with the stochastic solutions obtained from 50 simulations.

In Table 6.3 the results of the test are shown. Specifically, we refer with the term "fixed first-stage to deterministic" solution (FFS2D) to the solution obtained by solving the stochastic problem fixing the first-stage variables to the optimal sizing obtained by solving the deterministic model (D). Similarly, we refer with the term "fixed first-stage to stochastic" solution (FFS2S) to the solution obtained by solving the stochastic problem fixing the first-stage variables to the optimal sizing obtained by solving the stochastic model (S), with 100 scenarios. Moreover, we

will evaluate how the results vary when fixing the first-stage variables, i.e., we will compare the KPIs between the solutions FFS2D and D, and the solutions FFS2S and S, in order to asses the stability of the models to perturbations of the inputs. Finally, we will analyze the differences between the solutions FFS2D and FFS2S, to verify with model between the deterministic and the stochastic achieves a better performance on the newly generated scenario set.

	Solution		Difference		
KPI	FFS2D	FFS2S	FFS2D - FFS2S	D - FFS2D	S - FFS2S
NPV [€]	129319	188432	+46%	-27%	-4%
Payback time	12	10	-2	+2	+0
Self-sufficiency	39.41%	43.24%	+4%	-13%	-1%
Self-consumption	76.29%	65.03%	-9%	+10%	+1%
Shared energy	62.18%	54.25%	-8%	+7%	+1%

Table 6.3: Comparison between the stochastic and the deterministic solutions for Configuration 1 on a new test set of 1000 generated scenarios.

The performance of the FFS2S solution tested on the new scenario set remains similar to the stochastic solution one summarized in Table 6.1. This is because a sufficiently large number of Monte Carlo simulations was chosen to reach a desired level of confidence and accuracy, which keeps the solution stable when tested on a different scenario set. The opposite, instead, is true for the optimal sizing provided by the deterministic solution. When stochasticity is introduced in the input data, the evaluated economic and energetic KPIs of the FFS2D solution vary significantly compared to the deterministic solution. As matter of fact, we observe a decrease of approximately 27% in NPV and consequently and increase of 2 years in the packback time. Moreover, we observe also a significant difference in the energetic KPIs, when passing when fixing the sizing variables of the deterministic solution: the introduction of the uncertainty leads to a decrease of the 13% in the self-sufficiency, and to an increase of the 10 10% and of the 7% in the self-consumption and shared energy percentage respectively. We now compare the results of the solutions with the first stage fixed. We immediately notice a huge discrepancy in the NPV: the FFS2D solution has produces an objective function lower of the 47% compared to the FFS2S solution, resulting in a 2-year increase in the payback time of the investment. Regarding the energetic KPIs, the FFS2D solution shows a decrease in self-sufficiency and an increase in self-consumption and shared energy percentage compared to the FFS2S solution due to the total installed kW being less.

What we can conclude from this analysis is that solving the deterministic model,

i.e., by simply considering the average profiles of generation load and average prices, we are not able to capture effectively the variability of these parameters and how their interact among them. These results suggests that the stochastic solution is more stable to variations in the input data, and thus better to optimally size the REC under uncertainty.

Chapter 7

Conclusions

In this thesis we have presented optimization model to optimally size the PV generation plant and storage system of a Renewable Energy Community. After a literature analysis, the most appropriate methodology for the given task found was the development of a two-stage stochastic optimization problem, through which it has been possible to distinguish between the initial decision phase, i.e., regarding the sizing, from the second phase, which accounts for the energy dispatch operations. Three main uncertain parameters have been considered and modelled through a probabilistic approach. To consistently capture the variability in the PV production, a Gaussian Mixture distribution has been introduced to represent the generation in an hour of a given month. Sampling from these distributions, PV profiles are generated, which are then fed to the stochastic optimization model as inputs. Another critical parameter has been the energy load, and a distinction has been made between residential and non-residential buildings. In the study of the residential energy demand, social groups of consumers have been introduce in order to take into account the different possible patterns in the energy consumption curves. Finally, a time series analysis has been carried out for the introduction of the energy prices into the model, namely the PUN and bill price. On the basis of the historical data, three scenarios have been generated, representing high, average and low values for these prices. Having defined how the uncertainties are tackled and given the mathematical formulation of the model, a performance evaluation has been performed. The latter aimed at establishing the suitable number of scenarios to generate in order to achieve a desired trade-off between accuracy and computational time and at investigating which of the uncertain parameters have the greatest impact on the final output. In the end, three different configurations of RECs are simulated and tested, comparing the results among them. Moreover, the solution given by the stochastic model has been compared to the solution given by the deterministic problem, obtained by setting the uncertain parameters to their mean values. What we show is that the stochastic solution is more resilient

to perturbation in the input data and achieves a better performance in terms of the economic KPIs when tested on newly generated set of scenarios.

The developed model is open to further developments, due to its flexibility given by the separability between the uncertainty modelling and the problem formulation. Therefore, different possible improvements will be discussed. On the uncertainty modelling side, one future step might be to refine the PV production scenario generation. In our methodology the variables accounting for the PV production are independent. For a more coherent representation, it might be convenient to introduce a dependence between these variables and explore new sampling techniques. As to the optimization part, the formulation of the problem might be refined by passing from a two-stage to multi-stage stochastic approach. This switch introduces non-anticipativity constraints into the model, but harshly raises its complexity. Therefore, scenario reduction techniques might be essential to study and introduce for the running of such a model.

Bibliography

[1] Iea. *CO2 emissions in 2022 – analysis*. URL: https://www.iea.org/reports/co2-emissions-in-2022 (cit. on p. 1).

[2] Iea. *Greenhouse gas emissions from Energy Data Explorer – Data Tools*. URL: https://www.iea.org/data-and-statistics/data-tools/greenhouse-gas-emissions-from-energy-data-explorer (cit. on p. 2).

[3] European Commission. *Green Deal*. 2019. URL: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/delivering-european-green-deal_en (cit. on p. 6).

[4] European Commission. *Commissione Europea, Clean Energy for All European package*. 2019. URL: https://energy.ec.europa.eu/topics/energy-strategy/clean-energy-all-europeans-package_en (cit. on p. 6).

[5] European Commission. *Renewable Energy Directive II*. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2018.328.01.0082.01.ENG&toc=OJ%3AL%3A2018%3A328%3ATOC (cit. on p. 6).

[6] Presidenza del Consiglio dei Ministri. *Decreto-legge del 30/12/2019 n. 162. disposizioni urgenti in materia di proroga di termini legislativi, di organizzazione delle pubbliche amministrazioni, nonchè di innovazione tecnologica*. 2019 (cit. on p. 8).

[7] ARERA. *Deliberazione 4 agosto 2020 318/2020/r/eel regolazione delle partite economiche relative all'energia elettrica condivisa da un gruppo di autoconsumatori di energia rinnovabile che agiscono collettivamente in edifici e condomini oppure condivisa in una comunità di energia rinnovabile l'autorità di regolazione per energia reti e ambiente*. 2020 (cit. on p. 8).

[8] A Zakaria, Firas B Ismail, MS Hossain Lipu, and Mahammad Abdul Hannan. «Uncertainty models for stochastic optimization in renewable energy applications». In: *Renewable Energy* 145 (2020), pp. 1543–1571 (cit. on pp. 12, 14, 17).

[9] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. «A review of uncertainty characterisation approaches for the optimal design of distributed energy systems». In: *Renewable and Sustainable Energy Reviews* 88 (2018), pp. 258–277 (cit. on pp. 12, 18).

[10] Ayse Selin Kocaman, Emin Ozyoruk, Shantanu Taneja, and Vijay Modi. «A stochastic framework to evaluate the impact of agricultural load flexibility on the sizing of renewable energy systems». In: *Renewable Energy* 152 (2020), pp. 1067–1078 (cit. on pp. 12, 13, 17, 18).

[11] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. «Design of distributed energy systems under uncertainty: A two-stage stochastic programming approach». In: *Applied energy* 222 (2018), pp. 932–950 (cit. on pp. 12, 13, 15, 17–19).

[12] Stefano Moret. *Strategic energy planning under uncertainty*. Tech. rep. EPFL, 2017 (cit. on pp. 12, 13, 17–19).

[13] Arun Rathore and NP Patidar. «Optimal sizing and allocation of renewable based distribution generation with gravity energy storage considering stochastic nature using particle swarm optimization in radial distribution network». In: *Journal of Energy Storage* 35 (2021), p. 102282 (cit. on pp. 13, 15, 17–19).

[14] Masoud Sharafi and Tarek Y ElMekkawy. «Stochastic optimization of hybrid renewable energy systems using sampling average method». In: *Renewable and Sustainable Energy Reviews* 52 (2015), pp. 1668–1679 (cit. on pp. 13, 15, 17–19).

[15] Zhihui Zhang, Rui Jing, Jian Lin, Xiaonan Wang, Koen H van Dam, Meng Wang, Chao Meng, Shan Xie, and Yingru Zhao. «Combining agent-based residential demand modeling with design optimization for integrated energy systems planning and operation». In: *Applied Energy* 263 (2020), p. 114623 (cit. on pp. 13, 14, 17, 18).

[16] Zhe Zhou, Jianyun Zhang, Pei Liu, Zheng Li, Michael C Georgiadis, and Efstratios N Pistikopoulos. «A two-stage stochastic programming model for the optimal design of distributed energy systems». In: *Applied Energy* 103 (2013), pp. 135–144 (cit. on pp. 13, 15, 17–19, 40).

[17] Jiah Yu, Jun-Hyung Ryu, and In-beum Lee. «A stochastic optimization approach to the design and operation planning of a hybrid renewable energy system». In: *Applied energy* 247 (2019), pp. 212–220 (cit. on pp. 12, 13, 17, 18).

[18] George Cristian Lazaroiu, Virgil Dumbrava, Georgiana Balaban, Michela Longo, and Dario Zaninelli. «Stochastic optimization of microgrids with renewable and storage energy systems». In: *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*. IEEE. 2016, pp. 1–5 (cit. on pp. 12, 13, 17–19).

[19] Jiehui Zheng, Yanni Kou, Mengshi Li, and Qinghua Wu. «Stochastic optimization of cost-risk for integrated energy system considering wind and solar power correlated». In: *Journal of Modern Power Systems and Clean Energy* 7.6 (2019), pp. 1472–1483 (cit. on pp. 13, 15, 17).

[20] Thomas TD Tran and Amanda D Smith. «Stochastic optimization for integration of renewable energy technologies in district energy systems for cost-effective use». In: *Energies* 12.3 (2019), p. 533 (cit. on pp. 13, 15, 17–19).

[21] Mohsen Bashir and Javad Sadeh. «Optimal sizing of hybrid wind/photovoltaic/battery considering the uncertainty of wind and photovoltaic power using Monte Carlo». In: *2012 11th International Conference on Environment and Electrical Engineering*. IEEE. 2012, pp. 1081–1086 (cit. on pp. 13, 15, 17, 18, 51).

[22] S Surender Reddy et al. «Optimization of renewable energy resources in hybrid energy systems». In: *J. Green Eng* 7.1 (2017), pp. 43–60 (cit. on pp. 13, 15, 17).

[23] RalphB D'Agostino. *Goodness-of-fit-techniques*. Routledge, 2017 (cit. on p. 21).

[24] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021 (cit. on p. 23).

[25] John R Birge and Francois Louveaux. *Introduction to stochastic programming*. Springer Science & Business Media, 2011 (cit. on p. 23).

[26] Daniele Salvatore Schiera, Francesco Demetrio Minuto, Lorenzo Bottaccioli, Romano Borchiellini, and Andrea Lanzini. «Analysis of rooftop photovoltaics diffusion in energy community buildings by a novel GIS-and agent-based modeling co-simulation platform». In: *Ieee Access* 7 (2019), pp. 93404–93432 (cit. on p. 37).

[27] URL: `https://www.terna.it/it/sistema-elettrico/statistiche/pubblicazioni-statistiche` (cit. on p. 38).

[28] ARERA. *ARERA - Dati e statistiche*. URL: `https://www.arera.it/dati-e-statistiche` (cit. on p. 40).

[29] Martin Bicher, Matthias Wastian, Dominik Brunmeir, and Niki Popper. «Review on Monte Carlo simulation stopping rules: how many samples are really enough?» In: *Simul. Notes Eur.* 32.1 (2022), pp. 1–8.

[30] European Commission and Directorate-General for Energy. *Clean energy for all Europeans*. Publications Office, 2019. DOI: doi/10.2833/9937.

[31] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok. «A Novel Ultrathin Elevated Channel Low-temperature Poly-Si TFT». In: 20 (Nov. 1999), pp. 569–571.