



Increasing renewable energy penetration allows mitigating greenhouse gas emissions but implies dealing with more variability challenging the power system. The present research proposes two planning models to efficiently integrate renewable and storage units in both large and small-scale electric energy system are proposed. Solutions include increasing reserves, supplied by dispatchable units, or storage capacity, which provides flexibility and is a climate friendly solution.

European gas supply strongly depends on import from foreign countries. The problem of renegotiation of long-term natural gas contracts is tackled by estimating the dependence risk and the optimal resource allocation of the underlying assets through pair-vine copulas and portfolio optimization methods. An optimization problem that describes the equilibrium state of a gas supply chain is instead developed to analyze the security of the Italian supplies.

LUIGI BOFFINO is currently a Power Intraday Trader in the private sector. Previously held the position of Power & Gas Data Analyst and Forecasting Specialist. He holds a PhD in Analytic for Economics and Business from the University of Bergamo (XXXI Cycle). His research focuses on Mathematical modeling for Electricity and Gas Markets, Energy Economics and computational methods for economics and finance, including stochastic programming, robust optimization, and pair copula modeling. He is member of AMASES (Italian Mathematical Society for Social and Economical Sciences), AIRO (Italian Association of Operations Research) and MOS (Mathematical Optimization Society).



Luigi Boffino

MODELS AND METHODS FOR ELECTRICITY AND GAS MARKETS IN A LOW-CARBON ECONOMY



**UNIVERSITÀ
DEGLI STUDI
DI BERGAMO**

Collana della Scuola di Alta Formazione Dottorale

- 23 -

Collana della Scuola di Alta Formazione Dottorale

Diretta da Paolo Cesaretti

Ogni volume è sottoposto a *blind peer review*.

ISSN: 2611-9927

Sito web: <https://aisberg.unibg.it/handle/10446/130100>

Luigi Boffino

**MODELS AND METHODS FOR ELECTRICITY AND GAS MARKETS
IN A LOW-CARBON ECONOMY**



Università degli Studi di Bergamo

2021

Models and methods for electricity and gas markets in a low-carbon economy /

Luigi Boffino. – Bergamo :

Università degli Studi di Bergamo, 2021.

(Collana della Scuola di Alta Formazione Dottorale; 23)

ISBN: 978-88-97413-40-0

DOI: [10.6092/978-88-97413-40-0](https://doi.org/10.6092/978-88-97413-40-0)

Questo volume è rilasciato sotto licenza Creative Commons

Attribuzione - Non commerciale - Non opere derivate 4.0



© 2021 Luigi Boffino

Progetto grafico: Servizi Editoriali – Università degli Studi di Bergamo

© 2018 Università degli Studi di Bergamo

via Salvecchio, 19

24129 Bergamo

Cod. Fiscale 80004350163

P. IVA 01612800167

<https://aisberg.unibg.it/handle/10446/175481>

Contents

1	Introduction	1
1.1	Electricity Market	2
1.1.1	History	2
1.1.2	Electric Power System Structure	2
1.1.3	Market organization	3
1.1.4	Participants	9
1.1.5	Networks	11
1.2	Gas Market	13
1.2.1	History	13
1.2.2	Gas Industry structure	14
1.2.3	Market Organization	16
1.2.4	Participants	19
1.3	Toward Decarbonization	19
1.3.1	Electricity role in a low-carbon energy economy	22
1.3.2	Gas role in a low-carbon economy	28
1.4	Thesis Motivations and Objectives	30
1.5	Methodologies	33
1.5.1	Optimization Problems	33
1.5.2	Stochastic Programming	35
1.5.3	Adaptive Robust Optimization	37
1.5.4	Pair-Copulas Construction	38
1.6	Thesis Organization	39

2	A Two-Stage Stochastic Optimization Planning Framework to Deeply Decarbonize Electric Power Systems	41
2.1	Introduction	42
2.2	Model	45
2.2.1	Model Notation	45
2.2.2	Model Formulation	49
2.3	Case-Study Data	55
2.4	Base-Case Results	62
2.5	Sensitivity Analyses	67
2.5.1	Technology Improvement	67
2.5.2	Transmission Congestion	70
2.5.3	Land Use	72
2.5.4	PV solar and BES	74
2.5.5	Adiabatic CAES	75
2.5.6	Electricity demand growth	77
2.6	Discussion and Conclusions	78
	Appendix	83
2.A	Data For Base Case	83
2.B	Data For Sensitivity Cases	87
2.B.1	Data For Technology-Improvement Sensitivity Case	87
2.B.2	Data For Land-Use Sensitivity Case	88
2.B.3	Data For Adiabatic-CAES Sensitivity Case	89
2.B.4	Data For PV and BES Sensitivity Case	91
3	Robust expansion planning of a distribution system with electric vehicles, storage and renewable units	93
3.1	Introduction	94
3.2	Problem Formulation	99
3.2.1	Notation	99
3.2.2	Planning horizon and short-term variability characterization	103
3.2.3	Deterministic model	104

CONTENTS

3.2.4	Long-term uncertainty characterization: uncertainty set	107
3.2.5	Adaptive robust optimization model	110
3.2.6	Definition of feasibility sets	111
3.3	Solution Procedure	113
3.3.1	Master problem	113
3.3.2	Subproblem	116
3.3.3	Algorithm	118
3.4	Case Study	119
3.5	Results	123
3.5.1	Impact of the long-term uncertainty	124
3.5.2	Impact of the revenues associated with EVs	127
3.5.3	Impact of the degree of dependency of distribution system from the main grid	129
3.5.4	Impact of the investment budget availability	130
3.5.5	An ex-post decarbonization analysis	131
3.5.6	Impact of the availability of a larger number of candidate nodes	134
3.5.7	Modeling long-term uncertainties using scenarios: a stochastic program- ming approach	138
3.6	Conclusions	141
Appendix		145
3.A	Additional data	145
4	Analysis of long-term natural gas contracts with vine copulas in optimization portfolio problems	149
4.1	Introduction	150
4.2	PCC approach and vine copulas	153
4.3	Data and dependence structure	156
4.3.1	Data analysis	156
4.3.2	Pair-Copula Constructions	158
4.4	The optimal composition of long-term natural gas contract	167
4.4.1	Optimization portfolio problems	167
4.4.2	Results	171

4.5	Conclusions	174
Appendix		175
4.A	Additional Figures	175
5	Evaluating the impacts of the external supply risk in a natural gas supply chain: the case of the Italian market	187
5.1	Introduction	188
5.2	The natural gas market	190
5.3	Modeling the gas supply chain with external supply risk	193
5.3.1	Notation	193
5.3.2	External supply risk indicators	196
5.3.3	Model assumptions	201
5.3.4	Optimization model for a natural gas supply chain with external supply risk	204
5.4	Case study	212
5.5	Results	217
5.5.1	No external supply risk	218
5.5.2	External supply risk	223
5.6	Conclusions	230
6	Closure	235
6.1	Summary	235
6.2	Conclusions	235
6.3	Future Research	236
List of Figures		238
List of Tables		242
Bibliography		245

Summary

Human activity is overloading the atmosphere with carbon dioxide and other greenhouse gas emissions, which trap heat and drive up the planet temperature, resulting in a negative impact on our health, environment, and climate. Governments are considering actions to curb climate change that will significantly change both electricity and gas markets. This thesis detects these issues and proposes models and methods to analyze how electricity and gas markets can contribute to the achievement of the decarbonization targets. Among the actions carried out to reduce carbon emissions, renewable energy penetration is the most effective one. However, the integration of wind and solar power plants in electric energy systems is extremely challenging because of the uncertainty and variability that characterize their electricity production. To accommodate the stochasticity of the renewable energy production, power systems need to be more flexible. This flexibility is provided by backup capacity in the form of reserves, which are provided by dispatchable units such as thermal plants, or batteries and storage devices, which represent an environmentally friendly solution.

Considering this framework, in the first part of this thesis, two expansion planning models to efficiently integrate renewable power plants, storage units, and electric vehicles in electric energy systems are proposed. We first want to detect which are the investment choices that have to be taken by a Market Operator to deeply decarbonize electric power system. To this aim, we propose a two-stage stochastic programming model to determine the optimal mix of generation and transmission capacity to build, taking into account both technical constraints and climate-related considerations. The model uses a mix of ac and high-voltage dc transmission lines, conventional and renewable generation, and energy-storage units to meet these objectives. Short- and long-term uncertainties are modeled using operating conditions and scenarios, respectively. Secondly, we take the view of a Distribution System Operator and we propose a stochastic adaptive robust optimization approach for the expansion of a small size electricity system problem. This involves

the construction of candidate renewable generating units, storage units, and charging stations for electric vehicles. In this case, long-term uncertainty is modeled using confidence bounds, while short-term uncertainty is represented through a number of operating conditions.

Gas-fired power plants represent the energy choice that can help to achieve a secure, competitive, and decarbonized power systems since they can significantly contribute to emission reduction by replacing high carbon fuels in electricity generation. In addition, these units are the ideal partner for variable renewable energy, providing back up to wind and solar. In the last years, Europe has taken the lead in the decarbonization policies and has imposed a strict carbon reduction target that has to be achieved by 2050. For all these reasons, in the second part of the thesis, we focus our attention on the European gas market and we detect two important issues that can affect its stability. The first one regards the re-negotiation of the long-term gas contracts invoked by European mid-streamers and the second one concerns the security of external supply. The need of re-negotiation arises from the fact that, in Europe, gas is sold according to two main methods: oil-indexed long-term contract and hub pricing. The fall of the European gas demand combined with the increase of the US shale gas exports and the rise of liquefied natural gas availability on international markets have led to a reduction of the European gas hub prices. Since oil-indexed long-term contracts have failed to promptly adjust their positions, European gas mid-streamers asked for a re-negotiation of their existing contracts to obtain new contracts linked to hub spot prices. In this thesis, we tackle this problem by estimating the dependence risk and the optimal resource allocation of the underlying assets of a gas long-term contract through pair-vine copulas and portfolio optimization methods, using different risk measures. We also investigate the risk of external supply because Europe mainly relies on imports from economically or politically unstable countries to cover its gas demand since the local production is very limited. The analysis of the external supply risk is focused on the Italian gas market whose demand is covered by 90% by imports from foreign countries. An optimization problem that describes the equilibrium state of a gas supply chain, where producers, mid-streamers, and final consumers exchange natural gas and liquefied natural gas both with long-term contracts and on spot markets is developed for this purpose.

Chapter 1

Introduction

ENERGY is the milestone of modern society and its supply has a direct impact on the economic and social development of nations. Indeed, there is a very powerful link between economic growth and energy consumption. An optimal energy supply improves the quality of our lives, works and activities. In general, energy resources have not been in places where high consumption has developed, which implied the construction of reliable networks to transport it and stimulated cooperation between parties for energy supply. As a result energy management has now become a form of international political power. This thesis analyzes both the electricity and gas markets investigating their transformation needed to the meet carbon emission reduction goals adopted by 195 Countries in Paris in 2015, after the International Climate Conference.

In this chapter, we provide an introduction to the thesis work. First, we present an overview of the electricity market, which includes a description of the restructuring process of the electricity sector and of the structure of the electric power system, an explanation of the functioning of the principal electricity markets and a review of the most relevant market agents. Second, we present an overview of the gas market, providing a summary of the deregulation process, a description of the structure of the gas supply chain and its participants, and an explanation of the functioning of the principal gas markets. Third, we present an overview of the decarbonization process and we analyze the role played by electricity and gas in the achievement of a low-carbon economy; we also describe all the technologies used in the models developed in this thesis. Fourth, a general overview of the methodologies used in this dissertation is provided. Finally, we provide the thesis objectives and organization.

1.1 Electricity Market

1.1.1 History

The electricity market is the place where transactions involving electricity are conducted. Historically, because of its special characteristics, electricity has been regarded to be a public service. This approach justified, in the early days, a vertical integrated monopolistic structure of this sector.

In these “traditionally” regulated structures, a public institution or a single regulated company was in charge of the complete organization of the electricity supply chain, represented by generation, transmission, distribution and consumption of electric energy. Under this scheme, the relations between different electric utilities were generally characterized by voluntary cooperation in a number of areas, such as joint management of operating reserve capacity, exchange for emergency reasons and third-party use of grids for current transmission and distribution under terms negotiated by the parties concerned (see Gomez-Exposito et al. (2018)).

However, the lack of competitions often generated disparities in the cost of electricity between regions and caused technological stagnation besides influencing market transparency. For these reasons, in the 80’s several regions started their restructuring process toward liberalization (see Sioshansi and Pfaffenberger (2006)). Chile is recognized as a pioneer in this process because in 1982 separated the basic activities involved in the provision of electric power. Anyway, Fischer et al. (2000) argue that the Chilean system has involved less restructuring, less competition and more regulation than first meets the eyes since, for many years, a single holding company owned the largest generating company, the primary distribution company in the Santiago region and the primary transmission company serving the largest region of the country. Similar reforms were not introduced anywhere else until 1990, when the electricity industry was much more radically transformed in England and Wales which create the first electricity pool. Shortly thereafter, it was the turn of Argentina and Norway (1991) followed by many countries of Latin America and in Europe, as well as many States of the Northern America. In all these deregulation processes, the underlying idea was always the same: unbundling generation, transmission and retailing activities.

1.1.2 Electric Power System Structure

The structure of the electricity sector can be divided into four main components: **production**, **transmission** (high voltage network), **distribution** (low voltage network), and **consumption**

(see Gomez-Exposito et al. (2018); Von Meier (2006)).

Production centers generate electricity at voltage of several kilovolts ranging typically from 6 kV to 33 kV. These values allow finding a good compromise between conductor size and insulation. In fact, the first one depends on current level, while insulation depends on the voltage level of the device. The power is then immediately transformed to voltage of hundred of kilovolts (from 132 kV to 700 kV) to optimize long-distance transmission reducing the line losses. The **transmission** grid interconnects all the major production and consumption centers forming a very dense network which ensures high reliability also in the case of failure of few lines. Electric substations represent communication nodes and can lower or increase the voltage. Moreover they are also centers where system measurement, protection, interruption and dispatch equipment are sited. **Distribution** networks are composed of low voltage transmission lines, which bring electric power to consumers adapted to their needs (from 20.000 V to 220 V). The configuration of these grids is usually radial. **Consumers** connect to the voltage level that best suits their power needs, i.e. highly energy-intensive industries connect directly to the high voltage grid while small consumers are connected to the low voltage grid. For a similar principle, generating stations with a very small output can feed their electric power directly into the distribution network. Figure 1.1 provides a scheme of the electric power system structure.

1.1.3 Market organization

Generators, consumers and retailers can trade electricity on two main trading arenas: a **Pool** for short-term transactions and a **Futures market** for medium-term or long-term transactions (see Conejo et al. (2010)). To ensure the security of system operation and energy delivery the **Ancillary Services markets** are also needed. **Capacity market** is instead needed to provide sufficient reliable future capacity. A more detailed description of these mechanisms is provided below.

- **Pool**

The pool includes the *day-ahead market*, *intraday market* (not in the USA) and the *balancing market*. Most of the energy is traded on the *day-ahead market* which covers the 24 hour horizon before the real time operation. The *intraday market* generally operates after the closure of the day-ahead market and is cleared closer to delivery day. Because the forecasts on renewable sources production become more accurate when approaching to the real time operation, non-dispatchable units, mainly represented by solar and wind power plants (see

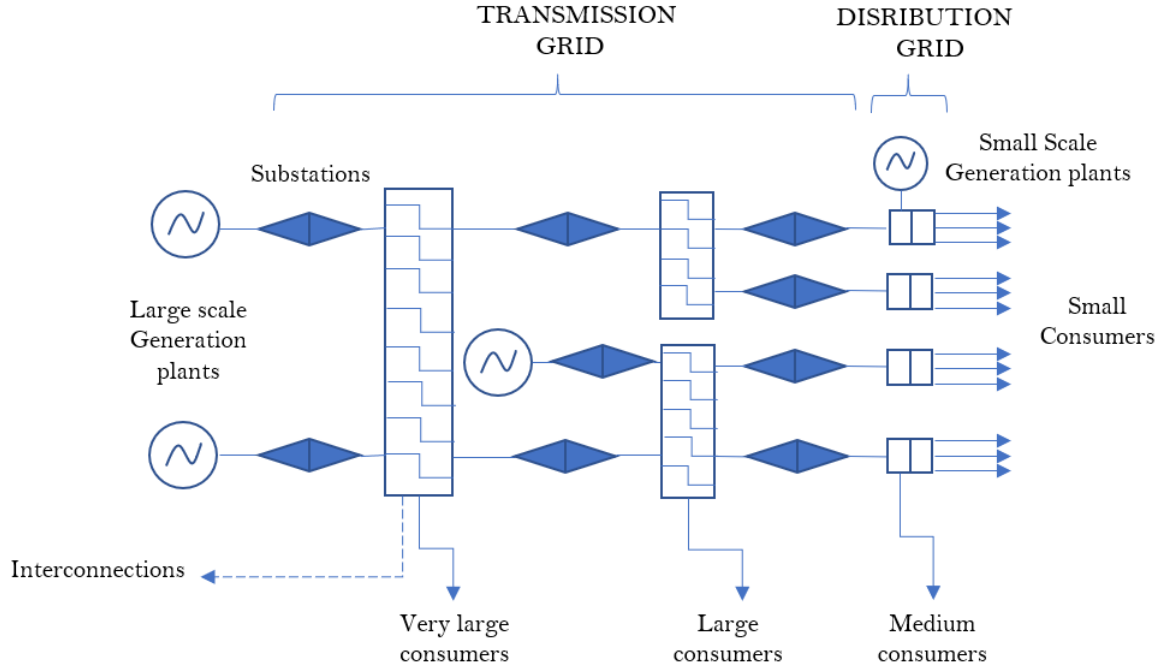


Figure 1.1: Electric Power System structure

Section 1.1.4), tend to rely more on the intraday market which have proven to be critical in accommodating large amounts of stochastic production (see Herrero et al. (2018)). The *balancing markets* are used to correct the real-time disparities that may arise between generation and demand due to equipment failures, load deviations or intermittent nature of some sources (e.g. wind or solar production). In the *day-ahead market* producers submit offers to the pool with their associated minimum selling price. At the same time, the retailers and consumers submit bids and their maximum buying price. The Market Operator collects purchase bid and sale offers and clears the day-ahead market using an appropriate market clearing procedure. The *day-ahead market* is cleared once a day, one day in advance and on an hourly basis. This procedure results in market clearing prices as well as production and consumption schedules. If the transmission grid is not considered in the market clearing procedure, then the resulting price is the same for all consumers and producers. Otherwise, if transmission network is taken into account, a location marginal price (LMP) is associated with each node of the power system. These prices may differ from each other due to line losses and/or congestions. The *intraday market* are either based on several successive auctions or on continuous trading

mechanisms.

The *balancing market* provides energy to cover both generation excess and deficit, settling deviations from the previous schedules.

- **Futures market**

In this market the energy is traded for delivery on a specified future date that range from one week or several years in advance. In order to do this, the participants buy or sell derivative products.¹ Thus, future markets are useful if the price of electricity is highly uncertain in the pool. This is the case of the electricity pool prices, which show high volatility, high percentage of outliers and other sources of uncertainty. Within this framework the future market represent a tool to hedge against the risk for both producers and consumers. On the future market, is it possible to subscribe different types of derivative products which includes, among the others, *bilateral contracts* and *options* (see Conejo et al. (2010)).

A *bilateral contract* is an agreement between two parties to exchange electric power under a set of specified conditions such as MW amount, time of delivery, duration, and price. Bilateral contracts can take the form of futures or forward contracts according to the market where they are traded (see El Khatib and Galiana (2007)). The former are generally traded in an exchange, and can be traded continuously up until their time of delivery. In contrast, forward contracts are typically negotiated in a Over The Counter (OTC) market directly between the load and generator with the terms of the contract remaining fixed until the time of delivery.

An *option* is an agreement for having the choice of delivering or consuming a specified amount of electricity in a future time period. In order to subscribe such agreement it is required to pay a premium, regardless of the eventual future deliver or consumption of energy.

- **Ancillary services markets**

The ancillary services markets provide all the services necessary for the operation of a transmission and distribution system. This markets are operated to guarantee reserve procurement and congestion management. The Network Codes² adopted in the European electricity market divide the ancillary services managed by the Transmission System Operator into *frequency* ancillary services, which deal with the balancing of the system and *non-frequency* ancillary services, which include voltage control and black-start capability. All these services are explained in the following:

¹A derivative is a financial security with a value that is reliant upon or derived from an underlying asset or group of assets. The derivative itself is a contract between two or more parties based upon the asset or assets. Its price is determined by fluctuations in the underlying asset. Hull and Basu (2016)

²See: https://www.entsoe.eu/network_codes/

- Frequency containment reserves (FCR) includes operating reserves necessary for constant containment of frequency deviations (fluctuations) from nominal value in order to constantly maintain the power balance in the whole synchronously interconnected system. This category typically includes operating reserves with the activation time up to 30 seconds. Operating reserves of this category are usually activated automatically and locally.
- Frequency restoration reserves (FRR) indicate active power reserves available to restore system frequency to the nominal frequency and to restore power balance to the scheduled value. This category includes operating reserves with an activation time typically between 30 seconds up to 15 minutes (depending on the specific requirements of the synchronous area). FRR replace FCR if the frequency deviation lasts longer than 30 seconds. FRR can be distinguished between reserves with automatic activation (aFRR) and reserves with manual activation (mFRR).
- Replacement reserves (RR) indicate the active power reserves available to restore or support the required level of frequency restoration reserve (FRR) to be prepared for additional system imbalances, including generation reserves. This category includes operating reserves with activation time from 15 minutes (in Continental Europe) up to hours.
- Voltage is controlled to operate the network within the voltage ranges and to maintain voltage stability. Voltage requirements are critical to secure planning and operation of a power system within a synchronous area. Reactive power helps the power grid maintain voltage levels. The reactive power can be defined as the difference in the phase between the voltage and current, or what additional voltage would be needed to restore the system to being in phase. Reactive power and real power (the power we actually consume) are substitutes in production. Markets Operators usually require generator to produce reactive power when requested, with any foregone real power consumption compensated based on opportunity cost.
- Black start is the capability of a power plant to start itself independent of the power grid. Some power plants have on-site generators that can be used to start the main turbine generator spinning. Others actually take electricity from the power grid for this purpose. Power plants that provide black start capability agree to keep an on-site generator or other grid-independent power source ready to operate. In the case of a large

blackout, these power plants will start themselves up using their on-site generation, and feed power into the grid so that other plants can start up and restore electricity service.

- **Capacity market**

The capacity market provides payments to power plants as an economic incentive to commit to being able to produce electricity at some point in the future, in addition to day-ahead/real-time energy and ancillary services. Capacity payments are designed to ensure that consumers continue to benefit from reliable electricity supplies at an affordable price. These payments can be received by existing power plants who promise to stay operational in the future, or by new power plants that commit to be finished with construction and operational by a certain date. Since capacity payments are received by generators regardless of whether the power plant actually produces any electricity, they represent a very attractive profit especially when market prices in the day-ahead and real-time energy markets are low.

1.1.4 Participants

The main agents participating in a restructured electricity market are introduced below (see Conejo et al. (2010); Shahidehpour et al. (2003)). They can be divided into non-institutional and institutional market agents. Among the non-institutional market agents we find:

- **Producers**

These agents own generation facilities that produce electricity or interact on behalf of plant owners with the markets. They are also called generating companies (GENCOs). Besides producing electricity they are also in charge of investments, operation, and maintenance of their generation facilities with the objective of maximizing their profits. It is possible to distinguish two different types of generating units owned by producers: dispatchable and non-dispatchable ones. *Dispatchable* units, mainly represented by conventional power plants such as coal and gas-fired units, provides reserves and regulatory capacity in addition to participating to the day-ahead and intraday markets. *Non-dispatchable* units are wind and solar power plants that, which due to their uncertain and intermittent production cannot actively participate to the ancillary services markets.

- **Consumers**

They are the end-users of the electricity. In general, large consumers can directly buy energy from the electricity markets, while small consumers mainly purchase electricity from the retailers. Their aim is to minimize their procurement costs or maximize the utility that they obtain from consuming electricity.

- **Retailers**

They are intermediaries who trade energy between producers and consumers. They buy electricity in the pool or in the futures markets and make profit by selling it to the consumers.

Institutional markets agents include:

- **Market Operator (MO)/Power Exchange (PX)**

It is the entity responsible for the economic management of the electricity market. It enforces the market rules and clears the markets, providing the final prices and the traded quantities.

- **Transmission System Operator (TSO)**

It is the entity in charge of maintenance and expansion of the transmission infrastructures used to transport the electric energy. It provides grid access to the electricity market players, according to non-discriminatory and transparent rules. Generally it also manages the reserve procurement and the balancing market.

- **Independent System Operator (ISO)**

It is the entity in charge of both the economical and technical management of the whole power system.

- **Market Regulator/National Regulatory Authority**

It is the entity that oversees the electricity market, ensuring its competitive and adequate functioning.

In non-restructured or traditionally regulated markets, as in some parts of the USA like Southeast, Southwest (except California) and Northwest, vertically integrated utilities are still responsible for the entire flow of the electricity to consumers.³ In other USA markets such as the PJM interconnection, the ISO-New England or the Californian CAISO, there is only a single entity, the ISO, who is in charge of both technical control and economic management of the system, while the transmission network is still owned by the vertically-integrated company. In Europe the economic management and the technical control of the power system are separated and managed by PXs and TSOs, respectively. For instance, considering the Italian electricity market GME (Gestore dei Mercati Energetici) acts as PX and Terna as TSO who is also in charge of managing the ancillary services markets (Mercato del Servizio di Dispacciamento). ARERA (Autorità di Regolazione per Energia Reti e Ambiente) is instead the Market Regulator (see Gestore Mercati Energetici (2009)).

³See: <https://www.epa.gov/greenpower/us-electricity-grid-markets>

1.1.5 Networks

In this thesis, we consider both transmission and distribution networks. As mentioned in Section 1.1.1, the main difference concerns the voltage, which is high in the first one while medium/low in the latter. Transmission lines can transport more electricity for longer distances with respect to distribution lines, which are shorter and transport electricity locally.

They present also a different topology. Transmission lines have usually a so-called “Network” configuration: in a network, any two points are usually connected by more than one path, meaning that some lines form loops within the system. This system is designed so that power can be injected at various locations and power can flow in different directions along the major transmission lines (bi-direction). This feature is important because since there are multiple paths for power to flow, if one transmission line is lost for any reason, all the load can still be served.

Distribution lines have usually radial configuration which means that lines branch out sequentially and power flows strictly in one direction. This property is crucial in the context of circuit protection, i.e. the interruption of circuits or isolation of sections in the event of a problem or fault. Circuit breakers can be located to isolate immediately upstream of the problem, interrupting service to all downstream components. Radial systems present also an economic advantage: smaller conductor sizes can be used toward the ends of the feeders, as the remaining load connected downstream diminishes (see Von Meier (2006)). Anyway, in the last few years Distributed Generation (DG) is being introduced on distribution networks. DG is a small-scale set of technologies (mostly renewable) dedicated to produce electricity close to final user. DG allows power loss reduction, distribution costs saving, CO₂ reduction and a decrease in the electricity prices. However, it adds production uncertainty and requires bidirectional flows and consequently new protection schemes (see Ackermann et al. (2001)).

Some differences also arise from a physical point of view. Transmission ac lines are based on a three phase supply system, i.e. three wires, mainly for an economic reason: these lines use fewer wires than common sense might suggest, and less conductor capacity than would be required to transmit an equivalent amount of power using only a single phase of ac. Distribution lines can be either three phases, in case of medium voltage or larger properties, or single phase for residential customers (see Von Meier (2006)).

Despite the fact that the transmission system is mostly ac, sometimes it is beneficial to use dc transmission. In this case, the power has to be converted both at the beginning and the tail end of the dc line through relatively costly solid-state devices called thyristors. Dc lines are used for very long distances: the technical reason is that dc eliminates the problem of a stability limit which poses a power transmission constraint on long lines with a significant inductance. Because its effect depends on ac frequency, inductance is irrelevant for dc, and therefore only the thermal limit (determined by the resistance of the line) applies. Moreover, above a certain line length which varies according to multiple conditions, dc lines become also more economically convenient than ac line. Besides transmitting bulk power over long distances, another application for dc lines in modern power systems is to provide an intertie between two ac systems that are not synchronous, so that power can be shared between these systems. In this thesis, we include both high voltage ac and dc transmission lines. The latter become particularly useful for connecting, also through undersea cables, large off-shore wind or solar facilities sometimes meant to represent long distance power exchanges among different States or countries.

In Europe, the TSO is responsible for building, maintaining and operating the transmission lines. Most of the European Member States have a unique TSO, which owns most of the transmission grid, but there are some exceptions. For instance, Germany has currently four different TSOs. In addition, from December 2008, the 43 different European TSOs from 36 European Countries signed in Prague a declaration of intent to create the ENTSO-E, i.e. European Network of Transmission System Operators for Electricity that is a supranational entity created to support the implementation of the EU energy and climate policy and to promote closer cooperation across Europe's TSOs. To this aim, the ENTSO-E has introduced, with guidance from the Agency for the Cooperation of Energy Regulators (ACER), the Network Codes for electricity, setting the rules for integrating renewable energy sources (RES) into the power system and creating a single and coordinated European electricity market.⁴ In the USA, it is instead possible to find different Regional Transmission Operators (RTOs).

The distribution network is managed by several Distribution System Operators (DSOs), although in some cases, one of more DSOs own the majority of the distribution network. DSOs receive the bulk energy from the transmission grid and distribute it to the consumers in different locations. As for the transmission network, in Europe it exists also a European Distribution System Operators' Association (EDSO) for Smart Grids, which comprises 36 leading European DSOs. The goal

⁴See: https://www.entsoe.eu/network_codes/

is to promote smart grids to achieve the EU's ambitious energy and climate objectives to 2020 and beyond.⁵

1.2 Gas Market

1.2.1 History

The natural gas industry involves activities like gas exploration and production, transportation, and delivery. All these activities require large investments and highly specific assets that once are constructed at some locations cannot be removed. Moreover, all the parties involved in these activities are interdependent. Such a complex structure exhibited scale economies, meaning that over a range of output, the per unit cost of their outputs declines as output increases. A single vertically integrated firm would have the lowest cost of production making the natural gas industry as a natural monopoly, as for the electric power industry. Due to its structure, the natural gas sector have been initially publicly owned or economically regulated by governments both in Europe and in the USA. This control has been introduced to prevent such vertically integrated firms to exercise market power to raise prices artificially (see Jess (1997)).

However, economic growth and technological improvements have reduced the importance of scale economies, lowered the cost of collecting large amounts of capital, and changed the perceptions about the potential for economic efficiency. Natural gas industry also faced a number of gas shortages and price irregularities among the different states into the 70's. As a consequence, in parallel to the electricity markets, the natural gas sector started in the 80's a restructuring process toward deregulation.

In the USA, the first step of deregulation was taken in 1978, through the legislation of the Natural Gas Policy Act (NGPA) which abolished some natural gas price ceilings. From that time on, it started a slow restructuring process culminating in the 1992 with the FERC Order 636, which introduced mandatory unbundling services in natural gas pipelines and other measures, establishing wholesale competition in the natural gas sector (see Lee (2004)).

In Europe, the deregulation process started in the early 90's. The European Commission issued in 1998 the Directive 90/30/EC concerning common rules for the internal market in natural gas. It was the first of three consecutive gas Directives imposing increasingly explicit and stringent rules

⁵See: <https://www.edsoforsmartgrids.eu/>

to allow the development of gas markets and hubs (see Correljé (2016)). Directive 2009/73/EC, also known as Third Package,⁶ unified the different national approaches, enhancing the EU-wide cooperation also through the European Network of Transmission System Operators for Gas (ENTSO-G), founded on January the first 2009, to achieve frictionless cross-border gas trade among European members. Moreover, to maintain a fully functioning and interconnected internal energy market, ENTSO-G, supported by the Agency for the Cooperation of Energy Regulators (ACER), propose a set of legally binding rules, known as gas Network Codes (see Karan and Kazdağlı (2011)).

1.2.2 Gas Industry structure

The natural gas industry consists of three main segments: *up-stream*, *mid-stream*, and *down-stream* (see Correljé (2016)).

In the *up-stream* segment the exploration and the production of gas takes place. Nowadays, natural gas exploration typically begins with geologists examining the surface structure of the earth, and determining areas where it is geologically likely that gas deposits, usually coinciding with oil deposit, might exist. Then, basic seismology is used to run additional tests. As the Earth's crust is composed of different layers, each with its own properties, energy in the form of seismic waves, traveling underground, interacts differently with each of these layers helping to locate underground fossil fuel formations. It is also possible to use magnetometers to measure small differences in the magnetic field and gravimeters, which measure the difference in the gravitational field. Anyway, the best way to gain a full understanding of subsurface geology and to discover the potential natural gas deposits in a given area is to drill an exploratory well. Since this option is expensive and time consuming, it is performed only when there is an high probability to find petroleum/gas formations.

Once gas is located, it is extracted with different procedures according to the composition of the rocks where it is trapped. If the deposit is found in permeable rocks, trapped below impermeable rock, gas can be extracted by drilling down through the impermeable rock into the permeable rock. In the case where gas and oil are trapped in the spaces within impermeable shale rock these has to be fractured to get the gas out with a process known as hydraulic fracturing, commonly called fracking (see Tissot and Welte (2012)). The fluid used to create micro-fractures in the shale is a mixture of water, sand, and additives pumped at high pressure (see Vengosh et al. (2013)).

⁶See: <https://ec.europa.eu/energy/en/topics/markets-and-consumers/market-legislation>

Once extracted the gas and oil are separated in the refineries, by distillation into fractions with different boiling points which are then further processed. Natural gas, as we use it, is almost entirely methane. However, when we find natural gas underground, it comes associated with a variety of other trace compounds and gases, as well as oil and water, which must be removed.

The *mid-stream* segment involves the transport of the gas to the local distribution grids and the large-scale industrial users and power plants. On a continental scale, generally, the gas is transported via high pressure transmission pipelines. Overseas the gas is transported in tankers as Liquefied Natural Gas (LNG).

A transmission pipeline is made of several components that ensure the efficiency and reliability of the system. Transmission pipes are made of strong carbon steel material and can measure from 12 to 120 centimeters in diameter, depending on their function. To ensure that the natural gas flowing through the pipeline remains pressurized, periodical compression is required along the pipe. This is accomplished by compressor stations, usually placed at 60 to 160 km intervals along the pipeline. Metering stations are also present along the pipe, measuring the gas flow. A great number of valves is present inside the pipe: they work like gateways and if open gas flows freely, otherwise they can be used to stop gas flow along a certain section of the pipe. Finally, centralized control stations collect data from the pipeline components to maintain safety and provide quick response in case of malfunctions, leaks, or any other unusual activity.

LNG is obtained by cooling down natural gas to -160° Celsius. This process, although expensive is useful for transportation because LNG volume is 600 times less than that of gaseous natural gas. Liquefaction has the advantage of removing oxygen, carbon dioxide, sulfur, and water from the natural gas, resulting in LNG that is almost pure methane. It is transported by specialized tanker with insulated walls, and is kept in liquid form by auto-refrigeration. Before usage, it needs to be vaporized to gaseous form through a process called regasification. Boil-off losses usually occur during transportation and storage (see Dobrota et al. (2013); Hasan et al. (2009)), in addition to the ones associated with liquefaction and regasification phases.

Transportation of natural gas is closely linked to its storage: if not immediately required, it can be put into storage facilities for later usage. Since the consumption of natural gas is seasonal, it is usually stored during summer when the request is lower and then withdrawn in winter to meet the higher demand. Gas is usually stored underground in large reservoir that can be: depleted gas reservoir, aquifers, or salt caves. Natural gas is injected into these formations, that have been previously reconditioned to create a sort of storage vessel, building up pressure as more natural gas is

added. The higher the pressure in the storage facility, the more readily gas may be extracted. Once the pressure drops to below that of the wellhead, there is no pressure differential left to push the natural gas out of the storage facility. This means that, in any underground storage facility, there is a certain amount of gas that may never be extracted. This is known as physically unrecoverable gas and it is permanently embedded in the formation. In addition to underground storage, gas can also be stored as LNG in appropriate tanks.

In the *down-stream* segment the local distribution grids deliver the gas to consumers. The distribution process is similar to the transportation one, but involves moving smaller volumes of gas at much lower pressures over shorter distances to a great number of individual users. Smaller-diameter pipes are used to transport natural gas to individual consumers from the so-called “city-gates”, which are the delivery points where the natural gas is transferred from a transmission pipeline. Distribution pipes are usually made of highly advanced plastic, because of the need for flexibility, versatility and the ease of replacement.

1.2.3 Market Organization

To obtain supplies of natural gas, energy companies, storage operators, traders, and other market participants involved in all or in parts of the gas chain use the wholesale markets. This is composed by two different trading arenas: the **Spot market** for short-term transaction and the **Futures market** for medium or long-term transactions. Outside of the wholesale markets, Long-Term Contracts on OTC markets are largely diffused.

Wholesale gas trading points are usually located at physical or virtual gas hubs, where buyers and sellers can trade in natural gas. While “citygates” are distribution points for a city, a physical hub is a major shipping, distribution, and delivery point, which may have significant storage capacity, connections with onshore and offshore gas pipelines, and infrastructures to off-take or load LNG cargoes. In a physical hub, such as the Henry Hub in the USA or Zeebrugge in Belgium, traders are required to bring the gas there and parties are required to book the same quantity for both entry and exit on a point-to-point transaction mode, assuming that pre-agreed gas injected into the network at one point will be taken off at a predetermined location.

On the other side, a virtual hub is defined through a pipeline grid (interconnected pipelines with no point of origin or end) representing the entire country or a trans-regional zone, managed by a TSO. All the gas within the virtual hub can be traded, irrespective of its actual physical location. In-

dividual buyers and sellers can book different quantities for entry and exit into the system without a predetermined destination, thus increasing the trading flexibility compared to a physical hub. For this reason, virtual hubs are generally more suitable for countries short of domestic natural gas supply and relying on various sources of supplies, including LNG imports. To be considered mature, an hub should fulfill five main requirements: liquidity, volatility, anonymity, transparency, and traded volumes. A mature hub is able to develop a reliable forward curve that can be used for risk management purposes. The final stage of maturity is when the hub is sufficient liquid to trade specific traded products such as indices pricing their physical transactions (see Heather (2015); Rademaekers et al. (2008)). In Europe, only the National Balancing Point (NBP) in the UK and the Dutch Title Transfer Facility (TTF) in the Netherlands can be considered mature with enough level of liquidity. The German hubs NetConnect Germany (NCG) and Gaspool are closer to NBP and TTF in terms of performance transactions, while all other hubs need to be further developed (see Agency for the Cooperation of Energy Regulators (2015)).

- **Spot Market**

The spot market generally includes a *day-ahead* gas market, an *intraday* gas market, a *locational* market and a regulated market for the *trading of stored gas*. This is also the design applied to the Italian gas market whose details are provided in the following.

On the *day-ahead* gas market, offers and bids related to the three gas-days after that of the trading session opening are selected. Transaction takes place in accordance with the continuous trading mode in which orders are executed according to price and time entered.

On the *intraday* gas market, offers and bids related to the gas-day corresponding to that of trading session opening are selected. Also here, transaction takes place in accordance with the continuous trading mode.

On the *locational* market, authorized users supply the gas TSO with quantities of gas needed to manage the physical demands within the balancing zones or deviations provided between overall injections and withdrawals on the network. The sessions of this market are held only upon request of the TSO and transactions take place in the form of auction trading.

On the market for *trading stored gas*, all bids and offers of gas stored are traded by authorized users and by the TSO. Transactions take place in the manner of auction trading.

- **Futures Market**

To minimize the risks inherent to spot trades, market participants sign *futures* contracts to buy and sell gas to be supplied in future months, quarters, seasons or years, at a price ne-

gotiated on the contract date. To make trading easier, these futures apply to standardized products, for example, the supply of 1 MWh for each gas day during the delivery period. Futures represent the average of the expected spot prices over a longer period, making them generally less volatile than spot products. They are used as a basis for determining the prices paid by end-consumers.

As in the electricity market, it is also possible to buy *options*, an agreement for having the choice of delivering or consuming a specified amount of gas in a future time period. In some markets the TSO subscribes long term options to buy/sell flexible gas quantities for balancing purposes (see European Network of Transmission System Operator for Gas (2012)).

- **OTC markets**

The most common bilateral contracts on OTC markets are *Long-Term Natural Gas Contracts* (LTCs), which commit producer and buyer over long durations, usually 20-25 years and were the most commonly used supply agreements in the gas sector, generally considered necessary to guarantee the security of supply. Historically, LTCs have been characterized by quantity and price clauses. The Take or Pay (TOP) quantity clause obligates the mid-streamer to take a certain quantity of natural gas or to pay for it. The price clause makes the LTC price linked to a bundle of oil-products (price indexation) to grant price stability to the producers and hence reduce their risk on investments. For this reason, gas infrastructures have been often developed taking into account the LTCs in place (see Abada et al. (2017)).

Considering the current organization of the gas market at international level, we can observe that LTCs dominate the Asian natural gas market, while the United States mainly trade natural gas on spot markets. Continental Europe is in the middle of a transition from LTCs to the spot pricing system, with a significant fraction of its gas supply still linked to LTCs. However, in Europe oil-indexed pricing had not reflected gas market fundamentals for some time, making the situation untenable. Mid-streamers asked for a re-negotiation of their contracts in order to link LTCs price to the spot gas prices. For this transition to be successful, there is the need of reliable markets prices which allow for both physical balancing and financial risk management on liquid trading hubs.

1.2.4 Participants

The main agents participating in a restructured gas supply chain are briefly introduced below and represented in Fig. 1.2 (see Holz et al. (2013); Egging (2013)).

- **Producers**

Producers extract and refine natural gas and sell it to traders trying to maximize their profits. These agents usually own production facilities and they are also in charge of new investments, besides operation and maintenance of active production facilities.

- **Traders**

Traders are in charge of transporting and selling natural gas to marketers. Transportation can be done via pipelines, or liquefaction, shipping and regasification infrastructures. They rent transportation capacity from the TSO to export the natural gas. They can also rent storage capacity from the Storage System Operator (SSO) to arbitrate between seasonal price variations. Their aim is to maximize their profit.

- **Transmission System Operators (TSO)**

The TSO manages the transportation network and rents out capacity to traders.

- **Storage system Operators (SSO)**

The SSO manages the storage capacity and rents our capacity to traders.

- **Marketers or mid-streamers**

Marketers buy gas from the traders to supply all types of consumers which include residential/commercial, industries and power companies. Their aim is to maximize their profit.

- **Consumers**

Consumers are the-end users of natural gas. They all buy gas from marketers. Their aim is to minimize their procurement costs or maximize the utility that they obtain from consuming gas.

1.3 Toward Decarbonization

Human activity is overloading the atmosphere with carbon dioxide and other greenhouse gas emissions, which trap heat and drive up the planet temperature, resulting in a negative impact on our

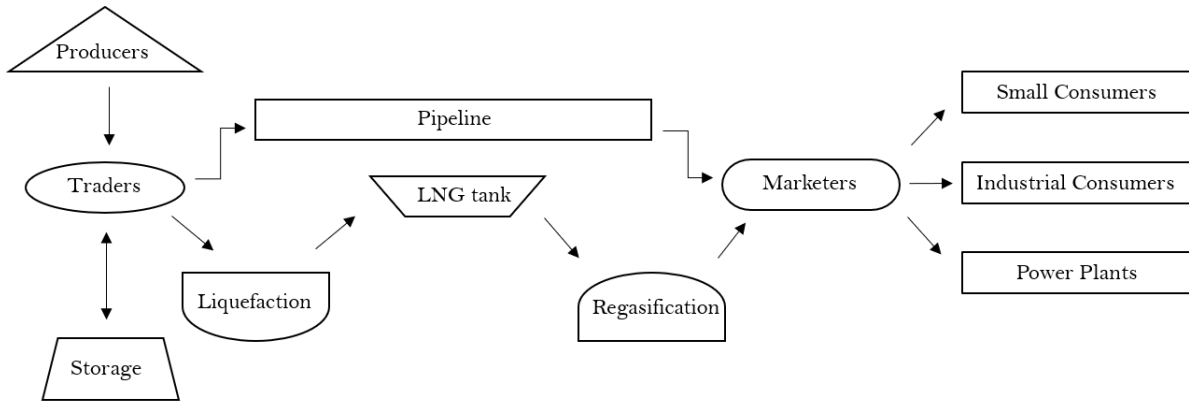


Figure 1.2: Representation of the gas supply chain participant's and interactions

health, environment, and climate. For these reasons, in December 2015, 195 countries signed the “Paris Climate Agreement”, which entered into force on November 4th, 2016. The targets of this agreement are ambitious and include a long-term goal of keeping the increase in global average temperature well below 2°C by 2100, preferably below 1.5°C and of undertaking rapid reductions of emissions in accordance with the best available technologies. However, taking into account the actual mitigation policies, the 1.5°C increase target has only 1% chance to be achieved and population growth is not a major contributing factor (see Raftery et al. (2017)). In fact, the long lifetime of CO_2 , the thermal inertia of oceans and others factors are responsible for some level of natural warming, which could be enough to exceed the Paris targets (see Mauritsen and Pincus (2017)). Consequences might be serious on different levels, including premature mortality from changes in air pollution attributable to climate change (see Silva et al. (2017)).

In addition to the Paris Agreement, a number of countries and regional international organizations have adopted carbon-reduction targets of their own.

The USA were among the first Parties to deposit their instruments of ratification, acceptance, approval or accession, allowing the Agreement to enter into force. However, on June 1 2017 they announced their withdrawal and asked for a re-negotiation.⁷ In spite of this, it is important to recall that according to the Article 28 of Paris Agreement the earliest date for a Party to leave the Agreement would be November 4th, 2020, upon expiry of one year from the date of receipt by the Depositary of the notification of withdrawal (see United Nation Framework Convention on Climate Change (2015)). Anyway, at the actual stage, the USA climate outlook is not favourable.

⁷See: <https://www.whitehouse.gov/briefings-statements/statement-president-trump-paris-climate-agreement/>

The Russian Federation instead, has not yet ratified the Paris Agreement, delaying the adoption of ambitious climate targets and policies until at least 2019.

In the rest of the world, China joined the Paris Agreement through the so-called Nationally Determined Contribution, committing to reach a CO₂ emission peak by 2030. Although under current policies, this target will most likely be achieved, it is still not compatible with the objective of limiting global temperature increase to below 2°C. On the other side, India is making an important effort and adopted in April 2018 a very ambitious plan called National Electricity Plan (NEP) to achieve the goal of a 40% non-fossil-based power capacity by 2030 at end of this year, i.e. 12 years earlier than targeted.

In the same line, Japan is currently developing its Basic Energy Plan and long-term energy strategy. There is uncertainty around the future role of nuclear, coal and renewable energy. Coal plant construction plans remain a concern and pose a serious risk to the government's future mitigation efforts.

Canada adopted the Pan-Canadian Framework on Clean Growth and Climate containing proposals for economy-wide measures, including a carbon pricing plan and a plan to phase-out traditional coal plants. Canada's Nationally Determined Contribution target is to reduce economy-wide GHG emissions by 30% below 2005 levels by 2030, but it remains unclear if it will rely on carbon sinks in forests, soils, and wetlands to achieve its target.

Finally, the European Union introduced in 2005 the European Union Emissions Trading System (EU-ETS) that is the first and largest cap-and-trade scheme applied on international scale to limit CO₂ emissions. The idea is the following: each ton of CO₂ emissions has to be covered by an allowance whose price is defined on a dedicated market. The EU-ETS went through two phases and it is now into the third one. The first phase, was a three year pilot (2005-2007) of "learning by doing" in preparation for the second phase, covering years 2008-2012. The over allocation of allowances that were assigned for free with a 95% and 90% proportions in first and second phases created some distortions reducing the direct impact of the EU-ETS on emissions. Since firms received the allowances for free they were able to generate the so-called "windfall profits" mainly in two ways: by selling the allowances on the market and by passing a significant part of the costs of CO₂ emission allowances through to product prices, resulting for example in higher electricity prices for consumers (see Sijm et al. (2006)). Despite this, the first two phases were able to generate saving on CO₂ emissions. Although disentangling the impact of the EU-ETS from other factors is complex, academic studies with both "top down", and sector-based "bottom up" evaluations attribute to this mechanism savings ranging from 2% to 4% of the total capped emissions per year

(see Laing et al. (2013)). The third phase, started in 2013, substantially modifies the EU-ETS trying to overcome these drawbacks by applying a single EU-wide emission cap in place of the previous system of national caps, by enlarging the involved sectors and by significantly reducing the amount of grandfathered allowances. This is also the tendency of the announced fourth phase that is currently under discussion. The European Union has also adopted the so-called Energy Roadmap 2050, committing to reduce greenhouse gas emissions by 80%-95% relative to 1990 levels by 2050. This implies a complete transformation of the energy system: improving energy efficiency, switching to renewable energy sources which should represent at least 55% of the gross final energy consumption in 2050 (actual level is around 10%), developing efficient and economically convenient storage mechanisms, enforcing the role of gas into the transition, developing smart technologies and alternative fuels including electric vehicles. Planning plays a central role in this transformation since large share of the current energy supply capacity will come to the end of their useful life in the next years. Moreover, there is the need of improving the condition for financing in the energy sector, decreasing the costs of capital for low-carbon investment. In this sense, carbon pricing can provide an incentive although increasing the risk of carbon leakage.⁸ A global effort is required and both electricity and gas market will face important transformations.

1.3.1 Electricity role in a low-carbon energy economy

Electricity will have to play a role much greater than now in order to contribute to decarbonization. According to the EU Roadmap the biggest share energy of supply technology in 2050 should come from renewables. Renewable Energy Sources (RES) have to be employed both on large and small scale, become more efficient and create economies of scale. Many renewable based technologies such as Concentrated Solar Power (CSP) or biofuels, need further development to reduce their costs, others such as wind plants and photovoltaic (PV) might be partially improved, for example by increasing the size of offshore wind turbines and blades to capture more wind or to improve the PV to harvest more solar power. Moreover, although renewable generation present several advantages such as unlimited resource availability all over the planet, very low production costs besides being environmentally sustainable, the operation of a renewable dominated electricity market it is not so trivial. This is due to the stochasticity that characterizes RES that makes uncertain the electricity production of such a kind of power plants (see Morales et al. (2013)). This have an effect on several

⁸Carbon leakage refers to a hypothetical situation where companies transfer production to countries with weaker climate policies in order to lower their costs.

floors:

- **Predictability:** first of all, when dealing with RES power production, forecasts become essential for an adequate integration of renewable power generation in electricity markets operations, since markets need be cleared in advance, while market participants shall make decisions even before that (see Section 1.1.3). An appropriate forecasting in a well-defined decision-making problem helps improving the decisions to be made, and allows controlling the risk generated by unforeseen events. Predictability has an impact on the cleared energy volumes and system prices, as well as their dynamics (see Jónsson et al. (2010)). This happens because the energy produced by RES units is commonly bid into the markets using forecasts of the future production, especially in the day-ahead and intraday markets. Indeed the potential need for balancing comes from the difference between day-ahead supply offers and actual generation. Consequently, it is not the actual renewable energy generation that drives the balancing direction, prices and quantities, but the forecasting errors instead. Moreover, since the marginal cost of renewable energy is very low, according to the so called merit-order principle, their production is scheduled before conventional production and their output directly influences the market price. As a result, in periods with high renewable power production the market price is low, while periods with low renewable power production are characterized by higher prices.
- **Variability:** renewable generation is intermittent and it is nonlinear function of the associated atmospheric variables such as solar irradiance or wind speed. Variability from one hour to the next, but also from one day to the other, is high and these patterns contribute to shaping the evolution of the various variables in day-ahead and balancing markets. Usually, variability in renewable production is not correlated to the fluctuation of the demand and so cannot be passively absorbed by consumers. For this reason, a high degree of variability in a power system requires a high degree of flexibility that can be achieved in several ways.
- **Flexibility need:** flexibility can be achieved by operating conventional units at production levels higher or lower than their optimum in an attempt to accommodate the variability of renewable generation by ramping up or down, which sometimes implies also starting-up or shutting-down conventional units. Consequently, if large variations of renewable generation have to be accommodated by conventional generation, this may result in conventional power plants operating in a less efficient way, thus reducing the emissions savings generated with

the aim of renewable production.

Another solution to increase flexibility is the implementation of direct and indirect control initiatives.

Direct control includes initiatives aimed at granting the TSO the right to directly modulate the demand by rationing or disconnecting single or groups of consumers. Typically, consumers involved in these strategies are protected by a contract fixing how often they can be disconnected or rationed.

Indirect control implies the use of economic incentives so that demand for power, rather than the supply, adapts to the stochastic and variable production. In practice, this is done by broadcasting time-varying prices to the consumers (dynamic real-time pricing). This strategy is also known as demand response and requires constant communication with consumer, achievable only with an upgrade of the grid, usually associated with the notion of smart grid.

A third way to increase flexibility is by optimally managing the reserves. Indeed, although part of the variability can be predicted, a renewable oriented power system increases the need for backup power to cope with unpredicted fluctuations of power production. This implies higher reserve availability which entails additional operating costs which can be minimized only by allocating and deploying reserves in an efficient manner.

Finally, probably the best way to increase flexibility and facilitate the integration of RES is by developing large-scale storage capabilities. Storage plants allow shifting in time the demand by storing electricity when demand is low (and so is the price) and releasing it, during high-demand (high-price) periods. From a practical point of view this translates into increasing the consumption during low-demand periods, allowing inflexible production units to maintain their power output or reducing the ramping of conventional units, while producing energy at high demand periods, reducing the need of reserves and the use of costly peaking power plants.⁹ Even though, currently, there exist a number of energy storage technologies that can contribute to this end, it seems that further development in this field is still needed for massive electricity storage to become a reality in the required magnitude (see Pickard et al. (2009)).

⁹These power plant generally run only when there is a high demand, known as peak load, for electricity. They receive a higher price for supplied power with respect to base load power plants which produce a consistent amount of electricity to meet the minimum demand (base load).

In the following the different types of renewable generation and storage facilities utilized in this thesis are briefly introduced.

Wind

A wind power plant is made of several wind turbines. Each turbine is supported by a conical steel tower which is firmly anchored to the ground. In the horizontal axis wind turbines, which are the most common worldwide and also the ones we consider in this thesis, the speed of the wind rotates the blades of a rotor, producing kinetic energy. The rotor then drives a generator, contained together with other mechanical components into the nacelle, that converts the mechanical energy into electricity (see International Energy Agency (2016c)). Wind plants can be built either onshore or offshore with the latter requiring both higher investment and operation costs with respect to onshore plants. Higher investments costs depends on more expensive marine foundations, installation procedures and integration in the electrical network, while higher operation cost are due to the limited access for operations and maintenance. Anyway, offshore plants experience also higher capacity factors¹⁰ with respect onshore plants because they can benefit from the higher wind speed, which generally increases with distance from the shore, and less turbulence, which allows the turbines to harvest the energy more effectively (see Bilgili et al. (2011)). Wind turbine prices have fallen by around half from 2009 leading to cheaper wind power globally. Moreover the global weighted average Levelized Cost Of Electricity¹¹ of onshore wind have dropped from USD 0.40/kWh in 1983 to USD 0.06/kWh in 2017, experiencing a 85% decline, with recent auctions in Brazil, Canada, Germany, India, Mexico and Morocco resulting in onshore wind power LCOEs as low as USD 0.03/kWh, making the technology one of the most competitive sources of new generation capacity (see International Renewable Energy Agency (2018b)). For these reasons, this technology is considered mature with respect to other more recent types of renewables. Large amounts of wind power have been integrated into power systems worldwide resulting in a total installed capacity of 5.4 GW at the end of 2017 (see Global Wind Energy Council (2017)). The projection of international agencies for future development of wind technologies estimate global wind installations reaching 743GW (of which 31GW offshore) by 2020 and 2,900-5,806 GW in 2050 according to different scenarios assumptions (see Global Wind Energy Council (2016)).

¹⁰The capacity factor is the ratio of the actual energy produced by a turbine in a period time, to the nameplate capacity of the turbine.

¹¹LCOE measures lifetime costs (including both investment and operations) divided by energy production. LCOE allows the comparison of different technologies of unequal life spans, project size, capital costs, risk, return, and capacities.

Solar Photovoltaic

A single Solar Photovoltaic device is known as a cell. Cells are connected together in chains to form larger units known as modules or panels which boost the power output of PV cells. Modules can be used individually, or several can be connected to form arrays. One or more arrays is then connected to the electrical grid as part of a complete PV system. Because of this modular structure, PV systems can be built to meet almost any electric power need, small or large. PV modules directly convert incident solar radiation into dc electricity, which can then be inverted to ac. PV gets its name from the process of converting light (photons) to electricity (voltage), which is called the PV effect. PV cells can be made of different materials: traditional solar cells are made from silicon and are usually flat-plate. Second-generation solar cells are called thin-film solar cells because they are made from amorphous silicon or nonsilicon materials such as cadmium telluride and are very thin and flexible. Third-generation solar cells are being made from a variety of new materials besides silicon, including solar inks using conventional printing press technologies, solar dyes, and conductive plastics. PV technology experienced a huge cost reduction in the last years with the module costs declined by 80% between the end of 2010 and the end of 2016. As a result also the weighted average LCOE cost experienced a 73% decrease to USD 0.10/kWh for new projects commissioned in 2017 (see International Renewable Energy Agency (2018b)). At the end of 2016 the worldwide installed cumulative PV capacity was 291GW. The projection of international agencies for future development of the technology estimates the global PV installations reaching 650 GW by 2020 and 3,155-7,122 GW by 2050 according to different scenarios assumptions (see International Renewable Energy Agency (2018a); International Energy Agency (2014a)).

Solar Thermal Electric Technology

Solar thermal electric technology, also known as Concentrated Solar Power with thermal energy storage (TES) relies on concentrating the sun's rays through the use of mirrors to create high temperature heat to drive a steam turbine. In the majority of today's systems, the sun energy is transferred to a fluid, which in turn is passed through heat exchangers to run a traditional electricity steam cycle, similar to the one used in conventional thermal power plants. Usually, a molten salt solution is the basis for generating steam because the heat in the fluid can be stored for a period of time into specific tanks and used later to generate electricity. According to the way used to collect the heat and generate electricity it is possible to distinguish various CSP configurations. These include parabolic trough collectors, solar power tower, linear Fresnel and Dish-Stirling (see Zhang

et al. (2013)). Solar power towers are currently the focal-point system currently most deployed, and also the one adopted in this thesis. CSP technology is still in its infancy in terms of deployment, although the cost have fallen between 2010 and 2017 and the estimated learning rate in the time horizon 2010-2020 is around 30%. Both investments and operating costs are expected to further decrease toward 2050. Auction results in 2016 and 2017 for projects that will be commissioned in 2020 and beyond, signed a step-change with the weighted average LCOE costs falling to between USD 0.06 and USD 0.10/kWh (see International Renewable Energy Agency (2018b)). At the end of 2016 the worldwide installed cumulative CSP capacity was 5 GW. The projection of international agencies for future development of the technology estimates the global CSP installations reaching 12 GW by 2020 and 633-982 GW by 2050 according to different scenarios assumptions (see International Renewable Energy Agency (2018a); International Energy Agency (2014b)).

Storage system: Batteries

Battery energy storage technologies, with their high energy densities, maturity of technology and relative ease of use represent a valuable option for small-scale energy storage. The batteries are made of stacked cells wherein chemical energy is converted to electrical energy and viceversa. The desired battery voltage as well as current levels are obtained by electrically connecting the cells in series and parallel. The batteries are rated in terms of their energy and power capacities, but they present also some other important features which differ from type to type. These are the efficiency, the life span (stated in terms of number of cycles), the operating temperature, the depth of discharge (batteries are generally not discharged completely), the rate of self-discharge and the energy density (see Divya and Østergaard (2009)). Battery storage is facing a significant development, and it is possible to distinguish many different types of batteries suitable for power system operations. These include lead-acid, sodium-sulphur, lithium-ion, nickel-cadmium, nickel-metal hybrid, and flow batteries. Lead-acid batteries are the oldest and most mature technology and have been largely implemented in power systems. The lithium-ion, sodium-sulphur, nickel-cadmium and nickel-metal hybrid batteries seem to represent the leading technologies in high-power-density battery applications. Of these, lithium-ion possesses the greatest potential for future development and optimization. In addition to small size and low weight the these batteries offer the highest energy density and a storage efficiency close to 100%, but they are cost expensive. The flow batteries are also promising for applications which require long duration storages due to their non-self-discharge capability, but they offer low power density (see Nair and Garimella (2010)). In

Chapter 3 we use batteries as small-scale storage solution within a small size electric energy system with electric vehicles and PV units.

Storage systems: Compressed Air Energy Storage

Compressed Air Energy Storage (CAES) represent a valid large-scale storage alternative for countries with limited water availability. At the time being, all the operating large-scale CAES plants are diabatic i.e. they allow energy exchanges with their surroundings. For this reason, diabatic CAES is considered a hybrid large-scale storage technology and uses both electricity (to compress air) and natural gas. The operation of a diabatic CAES plant can be divided into two phases. In the charging phase, that occurs during off-peak periods, CAES systems utilizes electrical energy from the grid to compress the air, that is stored at a high pressure and ambient temperature in a reservoir that can be an underground cavern, or any type of environment with similar features. The heat resultant from each compression is released in the atmosphere. In the discharging phase, a single-cycle gas turbine generator combines compressed air with natural gas in a combustion chamber. Combustion produces high-pressure gas, which is then expanded through a turbine, that drives both a generator and the input air compressor. The electricity produced is then delivered to the grid. In the last years another type of CAES, called adiabatic have been developed, although no commercial-scale operating systems exist today. Adiabatic CAES is an emission-free, pure storage technology which exploit the principle of storing the heat generated in the compression phase into dedicate thermal energy storage (TES). Then, during the discharging phase, the heat is released into the compressed air increasing its temperature with no need to burn gas and so avoiding CO₂ emissions. Additional details for these technologies, including investments and operation costs are given in Chapter 2.

1.3.2 Gas role in a low-carbon economy

In recent years technology advances such as horizontal drilling and hydraulic fracturing have unlocked gas resources in unconventional reservoirs, such as tight sands, coal bed methane, and shale rock rich in organic materials. The USA have been the forerunners of this new exploration process which is now applied internationally. As a result, unconventional gas should account for nearly half of the growth in global gas production to 2035, with China, USA and Australia as majors contributors. Consequently, gas resource estimates have increased sharply, and natural gas has now the potential to shed the supply, price volatility, and energy security concerns that have surrounded

it during the last few decades (see Flavin and Kitasei (2010)). In addition to that, natural gas is the only fossil fuel for which global demand is expected to grow, showing that it fares well under different policy conditions (see International Energy Agency (2016d, 2018)). The prospect of more abundant and economical gas supplies, together with the increasing urgency of the climate problem, is drawing more attention to the role that natural gas will play in the transition to a low-carbon power sector. More specifically, the role of gas will be critical for several reasons (see Holz et al. (2013)). In the short-medium term natural gas will substitute other fossil fuels with a relatively higher carbon content per generated energy, particularly coal. This could help reducing emissions with existing technologies until at least 2030-2035 according to the forecasts of the EU Roadmap 2050. This effect may be intensified by the advantages provided by gas-fired power plants, which present lower investment costs, are rather quickly built and relatively flexible. And it is thanks to this flexibility that gas-fired power plants can play a balancing role in an increasingly intermittent electricity system, acting as backup units. Other sectors than electricity generation, i.e. transportation and heating may be affected as well. Moreover the development of the Carbon Capture and Sequestration mechanism (CCS), which can cover a broad range of technologies, represent an important option to transform the production from natural gas plants in low-carbon production and pursuing decarbonization targets. CCS is currently considered to be technically feasible at commercial scale, with both large-scale and small-scale projects operating all over the world (see Gibbins and Chalmers (2008)). In the following we provide a brief description of this technology, which is also included as possible option in the long-term investment model developed in Chapter 2. Anyway, to keep its competitiveness as fuel for electricity generation, natural gas management has to be improved. First the gas market needs more integration, liquidity, storage capacity and diversity of supply sources. In particular, the latter is fundamental for non-producing countries which are constantly exposed to the risk of supply failure due to economical and political reasons. This aspect is analyzed in Chapter 5 of this thesis. In addition, LTCs may continue to be necessary to guarantee investments in both up-stream and mid-stream segments, but more flexibility in the price formula is required. There is the need of moving away from pure oil-indexation, opening to gas spot prices as possible underlying of these long-term contract. This aspect is developed in Chapter 4 of this thesis.

Carbon Capture and Sequestration

Carbon Capture and Sequestration is a technology that can capture up to 90% of the CO₂ emissions produced from the use of fossil fuels in electricity generation and industrial processes, preventing the carbon dioxide from entering the atmosphere. The CCS chain consists of three parts: capturing the carbon dioxide, transporting the carbon dioxide, and securely storing underground the carbon dioxide emissions in depleted oil and gas fields or deep saline aquifer formations. Capture technologies separate carbon dioxide from gases in electricity generation and may be done in at least three different ways: pre-combustion capture, post-combustion capture and oxy-fuel combustion. A pre-combustion system involves first converting solid, liquid or gaseous fuel into a mixture of hydrogen and carbon dioxide using processes such as gasification or reforming. The CO₂ is then captured. In the post-combustion CO₂ is captured from the exhaust of a combustion process in several ways which include absorbing it in a suitable solvent, using high pressure membrane filtration, with the aim of adsorption/desorption processes or cryogenic separation. In the process of oxy-fuel combustion the oxygen required is separated from air prior to combustion and the fuel is combusted in oxygen diluted with recycled flue-gas rather than by air. This oxygen-rich, nitrogen-free atmosphere results in final flue-gases consisting mainly of CO₂ and water, so producing a more concentrated CO₂ stream for easier purification. Fossil Fuel power plants can be built with the carbon capture technology integrated, or can be built “carbon capture ready”, which allows the plant to have carbon capture capabilities in the future. The cost of CCS is expected to decrease in the long-run as a result of research and development activities, and the building of economies of scale. Anyway also in this case, higher carbon prices are needed to push companies to invest in CCS (see Holz et al. (2013)).

1.4 Thesis Motivations and Objectives

The issues discussed in Section 1.3 has motivated the work carried out in this thesis, whose aim is to develop different modeling frameworks to represent the transition of electricity and gas markets toward a low-carbon economy. The thesis is divided into two main parts:

- In the first part, two planning models to efficiently integrate renewable and storage units in electric energy systems are proposed. Among the actions carried out to mitigate greenhouse gas emissions, increasing RES penetration is one of the most important. Anyway, renewable generation implies dealing with production variability and uncertainty, which represent a

challenge for the power system. Therefore, more backup capacity is needed. Possible solutions include increasing reserves, which are supplied by dispatchable units such as thermal plants. An alternative is the integration of storage capacity, which provides flexibility to system and represents an environmentally friendly solution. The objectives of this first part include:

- Developing an accurate modeling of multi-scale uncertainties, i.e. short-term uncertainty, able to represent daily variability of several factors including renewable energy production, and long-term uncertainty, which pertains to future projections. The problem is tackled by applying two different methodologies: stochastic programming and robust optimization.
- Integrating and adequately representing daily operations of different types of storage facilities, which help providing flexibility to the system.
- Correctly taking into account realistic network constraints by representing the topology of the system under study. Lines capacity availability is usually a concern because power production centers are located away from demand centers. We seek to develop models which allow lines expansion.

These topics are investigated in Chapters 2 and 3. The specific objectives of these two chapters include:

- To develop a two-stage stochastic programming investment model for generation- and transmission-capacity planning decisions taking into account decarbonization targets to be achieved by 2050 that is the considered reference year. Investment decisions are taken at the first stage, whereas operating decisions are taken at the second stage.
- To represent two types of climate policies: an explicit cap on carbon emission and a carbon price. In this way our model can represent capacity expansion under current European climate policy and policies that have been contemplated in the United States.
- To investigate how different technologies can contribute to the decarbonization of the energy sector. This includes the use of renewable technologies, such as wind and concentrating solar power (CSP), nuclear generation, and fossil-fueled generation with carbon capture and sequestration (CCS) systems. To allow both ac and high-voltage dc (HVDC) transmission lines, as well as energy-storage technologies to be built.

- To incorporate a novel representation of the operation of TES included into CSP plants in a long-term investment model.
 - To examine the impacts of a variety of policy and technical factors on power system decarbonization. Specifically, to study the effects of environmental policies, technology improvements, and transmission-network congestion. We also want to examine the trade-off between the use of wind and solar generation for decarbonization.
 - To develop an Adaptive Robust Optimization model for planning investments in a Small Size Electric Energy System (SSEES).
 - To model investments in PV power plants, storage units, charging stations for Electric Vehicles (EVs), and network lines in an integrate framework;
 - To study the changes in the investment decisions on the basis of the impact of long-term uncertainty, the central planner's revenues accruing from selling electricity to EVs at charging stations, the degree of autonomy of SSEES from the main grid, and the possibility of expanding the network.
 - To conduct an ex-post analysis on the amount of CO₂ saved with the utilization of EVs in order to evaluate the benefits of progressive road-transport decarbonization.
- In the second part, we analyze both the possible re-negotiation of the natural gas LTCs in the European market, characterized by an hybrid price system where also purchases of spot gas are allowed, and the security of supply of the Italian gas market. Gas is the less pollutant fossil fuel and will play an important role in the medium term, as explained in Section 1.3.2. A significant fraction of European gas supply depends on imports from foreign countries and, in addition, is linked to LTCs. The fall of the European natural gas demand, combined with the increase of the oil price, favored the emergence of a gas volume bubble that caused significant losses for most of the European mid-streamers because natural gas prices at the hubs were considerably lower with respect to the ones they were committed to pay because of LTCs. This situation instilled in the downstream part of the industry the idea of indexing contracts on gas spot prices rather than oil-based commodities. At this light also the issue of the security of supply of importing countries deserves attention. These topics are investigated in Chapters 4 and 5. Chapter specific objectives of this second part include:
 - To study via Pair-Copula Constructions (PCC), the dependence risk structure across the underlyings of LTCs on natural gas.

- To develop an integrated framework that combines vine copula and optimization techniques traditionally used in the contest of portfolio management.
- To address the debate over oil/spot indexation of European LTCs by defining the optimal composition of the assets traditionally used to price these contracts.
- To develop an optimization model that describes the equilibrium state of the natural gas supply chain where natural gas producers, mid-streamers, and consumers can sell and buy both natural gas and LNG through LTCs or/and on spot market.
- To study the risk of external gas supply by directly integrating new risk indicators in our model.
- To investigate different degrees of mid-streamers' flexibility by comparing a situation where they fully accomplish the LTC volume requirements with a case where they behave in a more flexible way.

1.5 Methodologies

The methodologies used in this dissertation are briefly revised below.

1.5.1 Optimization Problems

An optimization problem (OP) or mathematical programming problem consists in maximizing or minimizing a certain objective function subject to different restrictions, typically in the form of equality or inequality constraints. In general, an optimization problem can be formulated as follows (see Gabriel et al. (2013)):

$$\begin{array}{ll}
 \underset{x}{\text{minimize}} & f(x) \\
 \text{subject to} & h(x) = 0 \\
 & g(x) \leq 0
 \end{array}$$

where $x \in \mathbb{R}^n$ is the optimization variable vector, $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function to be minimized, $h(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{m_e}$ are the function for the equality constraints and $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{m_t}$ for the inequality constraints. The set of solution meeting both $h(t) = 0$ and $g(x) \leq 0$ constitutes

the feasible region. A solution within the feasible region is a feasible solution, while the feasible solution minimizing the objective function is the optimal solution.

The joint consideration of several interrelated optimization problems i where $i = 1, \dots, n$, constitutes an equilibrium problem of the form:

$$\begin{array}{ll} \underset{x^i}{\text{minimize}} & f^i(x^1, \dots, x^n) \\ \text{subject to} & h^i(x^1, \dots, x^n) = 0 \\ & g^i(x^i, \dots, x^n) \leq 0 \end{array}$$

where the vector $x^i \in \mathbb{R}^{n^i}$ includes the set of optimization variables of problem i . By setting $n_o = \sum_{i=1}^n n^i$ it is possible to define the objective functions and the constraints of the problem i as follows: $f^i : \mathbb{R}^{n_o} \rightarrow \mathbb{R}$, $h^i : \mathbb{R}^{n_o} \rightarrow \mathbb{R}^{m_e^i}$ and $g^i : \mathbb{R}^{n_o} \rightarrow \mathbb{R}^{m_t^i}$. The Karush-Kuhn-Tucker (KKT) conditions are a set of complementarity conditions which should be satisfied by the optimal solution of an optimization problem. KKT condition are a natural way to cast these types of equilibrium problems. A complementarity condition between a vector of non-negative variables $\mu \in \mathbb{R}^{m_t}$ and a function $g(x)$ can be defined as:

$$g(x) \leq 0 \quad \mu^\top g(x) = 0 \quad \mu \geq 0$$

which can be written more compactly as

$$0 \leq \mu \perp g(x) \leq 0$$

where $\perp$ indicates complementarity, i.e. $\mu^\top g(x) = 0$.

In order to meaningfully formulate KKT conditions associated to an optimization or equilibrium problem constraint qualification is required. KKT conditions can be necessary, but not sufficient conditions, i.e., solutions meeting them are not necessarily optimal but optimal solutions need to meet them. If the objective function of the related optimization problem is continuously differentiable and the set constituted by the constraints is convex, the solution is guaranteed to be optimal (sufficiency condition). KKT conditions are first-order conditions, i.e., conditions that are

formulated using first derivative vectors and matrices (gradients and Jacobians). The Langrangian function of the single optimization problem previously defined is:

$$\mathcal{L} = f(x) + \lambda^\top h(x) + \mu^\top g(x)$$

where $f(x), h(x), g(x)$ are assumed to be continuously differentiable in the feasible region. The associated KKT conditions are

$$\begin{aligned}\nabla_x f(x) + \lambda^\top \nabla_x h(x) + \mu^\top \nabla_x g(x) &= 0 \\ h(x) &= 0 \\ g(x) &\leq 0 \\ \mu^\top g(x) &= 0 \\ \mu &\geq 0\end{aligned}$$

where $\lambda \in \mathbb{R}^{m_e}$ and $\mu \in \mathbb{R}^{m_t}$ are Lagrange multiplier vectors associated with equality and inequality constraints, respectively, and ∇_x denotes the gradient with respect to the variable vector x . Note the the last three constraint can be expressed in compact form as stated above. In this dissertation we solve an equilibrium problem defined by jointly solving the KKT conditions of several single optimization problems each one modeling the behavior of a specific market player within the gas market.

1.5.2 Stochastic Programming

Most decision-making problems can be adequately formulated as optimization problems. If the input data of an optimization problem are well-defined and deterministic, its optimal solution is achieved by solving the problem. Anyway, when dealing with decision making problems in the real word, a common issue is the lack of perfect information. This happens in several fields, such as engineering, economics, finances, etc. However, decisions need to be made even with lack of perfect information. This is what motivates the use of stochastic programming models for decision making under uncertainty. Indeed, uncertain input data can be described through probability functions. In such a situation, possible solutions include substituting the uncertain input data by their corresponding expected values, which results in a well defined and deterministic optimization problem. However, solving such a problem may lead to a solution that, once implemented,

does not result in the best outcome. Alternatively, the probability distribution of input data can be approximated by a collection of plausible sets of input data with associated probabilities of occurrence. Doing so, it is possible to formulate a *stochastic* optimization problem to achieve a solution that is the best in some sense considering all possible values of the uncertain parameters.

The main drawback of stochastic programming is the size of the resulting problem, which usually leads to computational intractability. Moreover, as stated above, it is clear that stochastic programming relies on the knowledge of the probabilistic distribution of the uncertain parameters, which are usually represented with the aim of scenarios. Therefore adequate generation techniques are needed to achieve a good representation, and further, efficient scenario reduction techniques to attain computational tractability without losing important information (see Birge and Louveaux (1997); Kall et al. (1994); Prékopa (2013)). In this dissertation, a two-stage mixed-integer linear stochastic programming model is proposed. A general two-stage linear stochastic programming problem is formulated as follows:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x + \mathbb{E}_{\xi}\{Q(x, \xi)\} \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

where $Q(x, \xi)$ is the optimal value of the second-stage problem:

$$\begin{array}{ll} \underset{y}{\text{minimize}} & q(\xi)^T y \\ \text{subject to} & T(\xi)x + W(\xi)y = h(\xi) \\ & y \geq 0 \end{array}$$

This model can be applied to problems that deal with “here-and-now” decisions, which have to be taken on the basis of prior existing information about future situations without additional observations. The first-stage decisions do not depend on the scenario that will actually occur in the future. The second-stage decisions occur after the uncertainty realization and are the resources that can be used for each possible future scenario. The decisions taken depend on the future realization of the

scenario, for this reason are called “wait-and-see”. In this formulation $x \in \mathbb{R}^n$ is the first-stage decision variable vector, $y \in \mathbb{R}^n$ is the second-stage decision variable vector, and $\xi(q, T, W, h)$ contains the data of the second-stage problem. The “here-and-now” decision, x , is made at the first stage, before the realization of the uncertain data ξ is known. At the second stage, after the realization of ξ becomes available, an appropriate optimization problem is solved. At the first stage, the cost c^T , plus the expected cost of the optimal second-stage decision, $\mathbb{E}_\xi\{Q(x, \xi)\}$ are optimized. The second-stage problem can be viewed as an optimization problem which describes a supposedly optimal behavior when the uncertain data ξ are revealed. Matrices T and M are called technological and recourse matrices, respectively.

1.5.3 Adaptive Robust Optimization

Robust optimization is widely used to solve optimization problems under uncertainty. With respect to stochastic programming, it has the advantage of eliminating the need of scenario generation, and therefore reducing potential computational intractability. Indeed, instead of seeking to immunize the solution in some probabilistic sense to stochastic uncertainty, with robust optimization the decision-maker constructs a solution that is feasible for any realization of the uncertainty in a given set. The uncertain parameters are kept within robust sets, which are easier to construct with respect to scenarios. By considering the worst possible realization, the optimal solution is obtained, although this might sometimes lead to too conservative results (see Ben-Tal et al. (2009); Bertsimas et al. (2011)). The adaptive robust optimization formulation originates from the need of adequately represent a decision sequence in which (1) decision are made under uncertainty, (2) uncertainty realization is considered, (3) corrective actions can be made. A general adaptive robust optimization problem is formulates as follows:

$$\begin{aligned}
 \min_x \quad & \max_u \quad \min_y \quad f(\mathbf{x}, \mathbf{u}, \mathbf{y}) \\
 \text{s.t.} \quad & \text{s.t.} \\
 & \mathbf{h}^C(\mathbf{x}, \mathbf{u}, \mathbf{y}) = 0 \\
 & \mathbf{g}^C(\mathbf{x}, \mathbf{u}, \mathbf{y}) \leq 0 \\
 & \mathbf{y} \in \mathcal{Y} \\
 \text{s.t.} \quad &
 \end{aligned}$$

$$\begin{aligned}
 & \mathbf{u} \in \mathcal{U} \\
 \text{s.t.} \\
 & \mathbf{h}^I(\mathbf{x}) = 0 \\
 & \mathbf{g}^I(\mathbf{x}) \leq 0 \\
 & \mathbf{x} \in \mathcal{X}
 \end{aligned}$$

Objective function $f(\mathbf{x}, \mathbf{u}, \mathbf{y})$ represents the minimization of the total cost, including investment and operating costs. The investment decision variables are gathered in vector x , and depend on the so called expansion constraints. The entries of vector u are variables describing the uncertain parameters and $\mathcal{U}$ represent the uncertainty set which defines variability limits of uncertain parameters. Similarly, the entries of vector y are the operating decision variables. Note that $\mathcal{U}$ defines the feasibility set, which includes constraints related to the system operation and system limits. The worst case realization of the uncertainty and the successive adaptive actions are considered in the *max* – *min* right-hand-side problem, while the *min* left-hand-side problem seeks minimum total cost.

1.5.4 Pair-Copulas Construction

Copulas are functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions. Alternatively, copulas are multivariate distributions functions whose one-dimensional margins are uniform on the interval $[0,1]$. In a sense, every joint distribution function implicitly contains both a description of the marginal behaviour of individual variables and a description of their dependency structure. Copulas provide a way of isolating the description of their dependency structure. In particular, the copula approach allows capturing the complex dependency patterns of multivariate data, such as asymmetry and dependence in the extremes which cannot be captured by multivariate normal or student-t distributions.

The following theorem (Sklar) holds (see Sklar (1959)):

$$\begin{aligned}
 F(x_1, \dots, x_d) &= C(F_1(x_1), \dots, F_d(x_d)) \\
 f(x_1, \dots, x_d) &= c(F_1(x_1), \dots, F_d(x_d)) f_1(x_1) \dots f_d(x_d)
 \end{aligned}$$

where $F(x_1, \dots, x_d)$ and $f(x_1, \dots, x_d)$ are the joint distribution and density function, respectively. $F_i(x_i)$ and $f_i(x_i)$ for $i = 1, \dots, d$ are the marginal distribution and density function, respectively. $c(u_1, \dots, u_d) := \frac{\partial^d}{\partial u_1 \dots \partial u_d}$ is the copula density function.

While there exists a multitude of bivariate copulas, the class of multivariate copulas is still quite restricted. Hence, if the dependency structures of different pairs of variables in a multivariate problem are very different, not even the copula approach will allow for the construction of an appropriate model. It has been proved that under some regularity conditions a multivariate density can be expressed as a product of bivariate copulas, acting on several different conditional probability distributions. This has allowed formulating an approach finalized to the probabilistic construction of multivariate distribution functions based on pair-copulas, namely pair-copula construction (PCC). Many PCC's are feasible and, for this reason, there exist graphical structures to organize them. PCC allows exploiting the advantages of using copulas, providing additional flexibility by modeling conditional pairs.

1.6 Thesis Organization

This document is structured according to the following chapters:

Chapter 1 In this chapter we first present an overview of the electricity market, which includes a description of the restructuring process of the electricity sector and of the structure of the electric power system, an explanation of the functioning of the principal electricity markets and a review of the most relevant market agents. Second, we describe the gas market, providing a summary of its deregulation process, a description of the structure of the gas supply chain and of its participants, and an explanation of the functioning of the principal gas markets. Third, we present an overview of the decarbonization process and we analyze the role played by electricity and gas to meet low-carbon economy targets. We also describe all the technologies used in this thesis. Fourth, a general overview of the methodologies used in this dissertation is provided. Finally, we provide the thesis objectives.

Chapter 2 presents a two-stage stochastic optimization model to determine the optimal mix of generation and transmission capacity that needs to be built in order to serve future demands at least cost, while respecting both technical and environmental constraints. The model uses a mix of ac and high-voltage dc transmission lines, conventional and renewable

power plants, and energy-storage units to meet these objectives. Short- and long-term uncertainties are modeled using operating conditions and scenarios, respectively. A case study and a sensitivity analysis are provided to show how the proposed model works.

Chapter 3 proposes a stochastic adaptive robust optimization approach for the expansion planning problem of a small size electric energy system. This involves the construction of candidate renewable generating units, network lines, storage units, and charging stations for electric vehicles. Short- and long-term uncertainties are modeled using operating conditions and confidence bounds, respectively. A case study is used to illustrate the effectiveness of the proposed technique, while an ex-post decarbonization analysis is conducted to evaluate the environmental impact of the integration of electric vehicles in the energy system.

Chapter 4 originates from the European debate concerning oil versus gas indexation of long-term natural gas contracts. The dependence risk and the optimal resource allocation of the underlying assets of a gas long-term contract are estimated through pair-vine copulas and portfolio optimization methods with respect to different risk measures, respectively. Both spot gas prices traded at the hub and oil-based commodities are considered as possible underlyings of the LTCs. Results can provide concrete indications for practitioners in the field..

Chapter 5 analyzes the security of external supply of the Italian gas market that mainly relies on natural gas imports to cover its internal demand. An optimization problem, which describes the equilibrium state of a gas supply chain where producers, mid-streamers, and final consumers exchange natural gas and LNG, is developed. Both long-term contracts and spot pricing systems are considered. Mid-streamers are assumed to be exposed to the external supply risk, which is estimated with indicators that we develop starting from those already existing in the literature. In addition, different degrees of mid-streamers' flexibility are investigated by comparing a situation where mid-streamers fully satisfy the long-term contract volume clause to a case where the fulfillment of this volume clause is not mandatory.

Chapter 2

A Two-Stage Stochastic Optimization Planning Framework to Deeply Decarbonize Electric Power Systems

In 2015, 195 countries signed the Paris Agreement under the United Nations Framework Convention on Climate Change. To achieve the ambitious greenhouse gas-reduction targets therein, the electric power sector must be fundamentally transformed. To this end, we develop a two-stage stochastic optimization model. The proposed model determines the optimal mix of generation and transmission capacity to build to serve future demands at least cost, while respecting technical constraints and climate-related considerations. The model uses a mix of AC and High-Voltage DC transmission lines, conventional and renewable generation, and different types of energy-storage units to meet these objectives. Short- and long-term uncertainties are modeled using operating conditions and scenarios, respectively.

We demonstrate the model using a case study that is based on the Texas power system, with 2050 as the target year of the analysis. We include explicit carbon-emissions constraints. Doing so allows us to examine the effect of carbon-reduction targets and deep decarbonization of electricity production on investment decisions. As expected, we find that thermal-dominated power systems must transition toward having a renewable-dominated generation mix.

2.1 Introduction

As a result of concerns surrounding climate change, 195 countries signed the Paris Agreement under the United Nations Framework Convention on Climate Change in 2015. The agreement includes ambitious climate-related goals, which necessitate rapid and substantive declines in carbon emissions and the carbon-intensity of human activity. Moreover, a number of countries and regional international organizations have adopted carbon-reduction targets of their own. For instance, the European Union has adopted the so-called Energy Roadmap 2050, which includes targets to reduce greenhouse gas emissions by 80%–95% relative to 1990 levels by 2050.

Electricity production represents a major source of anthropogenic carbon emissions. Thus, the electric power system must play a substantive role in these reductions. Greater use of renewable energy sources, such as wind and solar, is one possible means of decarbonizing electricity production. However, the real-time availability of weather-dependent renewable energy sources is variable and uncertain. This gives rise to challenges in relying on such energy sources in place of dispatchable (e.g., thermal) generation. Spatial disaggregation of renewables and energy storage are possible means to mitigate these issues that are caused by resource variability and uncertainty. There are a number of works in the literature that examine decarbonization pathways for electric power systems and the role of renewables therein. Di Sbroiavacca et al. (2016) show how climate-related goals and carbon reductions could be achieved in the Argentinean electricity sector. Along the same lines, Calderón et al. (2016) examine decarbonization of the Colombian power system. They find that absent climate policy, CO₂ emissions will increase for the foreseeable future with no stabilization. They further show that carbon-reduction goals can be achieved with greater use of renewable resources. Qi et al. (2016) examine how the carbon-intensity of electricity production in China was reduced between 2005 and 2013. The aim of this analysis is to better inform future policy designs for further decarbonization of the Chinese power system.

Although these works demonstrate that renewables can play a significant role in decarbonizing electricity production, renewables create planning and operating challenges. As an example, Graf and Marcantonini (2017) conduct an empirical study of the operation of the Italian power system, using historical data from between 2005 and 2014. They show that renewables did displace thermal generation during this period, thereby reducing emissions. However, they also find that the variability in real-time renewable output increases cycling of thermal plants, reducing their operating efficiencies and slightly increasing their emissions rates. There is, nevertheless, a net emissions reduction from renewables during this period, as the former effect far outweighs the latter.

Expanding upon the findings of Graf and Marcantonini (2017), achieving deep decarbonization of electricity production requires major changes in the generation mix. One must go beyond simply adding more renewables to a power system. Instead, the balance of the system must become more flexible to accommodate the variability of real-time renewable availability. Furthermore, as discussed by Denholm et al. (2010), energy storage is expected to have a growing role as renewables constitute a larger portion of the generation mix. This is because energy storage provides significant flexibility in shifting renewable generation through time, thereby managing its variability.

This chapter adds to this existing literature by developing a modeling methodology to make generation- and transmission-planning decisions with decarbonization targets. Our model is formulated as a two-stage stochastic optimization problem wherein generation- and transmission-capacity planning decisions are made in the first stage. The second stage represents system operations during a set of representative operating days. These operating days implicitly capture short-term/small-scale uncertainties, such as load and renewable-supply variability. Long-term/large-scale uncertainties that are related to load growth are captured explicitly through scenarios that define the second stage of the stochastic optimization model.

Our model is formulated from the perspective of a central planner (*e.g.*, an integrated-resource planning process that is undertaken by an electric utility or a regulator). A number of jurisdictions have restructured electricity markets in place, meaning that this type of central-planning process does not occur. Nevertheless, our model is of value in such settings, as it can provide guidance on long-term planning decisions. Moreover, our modeling methodology provides insights as to how investments should be made by a risk-neutral agent that operates in a perfectly competitive restructured market. Despite our making these assumptions, agents are risk-averse in practice and markets for risk are incomplete and imperfect. de Maere d'Aertrycke et al. (2017) investigate the issue of incomplete risk trading and its impact on investments. They do this by analyzing stochastic equilibrium problems to quantify the impact of risk, market completeness, and contracting on generation investment while considering riskmitigation instruments, such as contracts for differences, reliability options, and forward capacity markets. They show that these instruments are quite effective complements to the energy-only market, especially when they are controlled and managed by a public regulator. While risk-related issues are important, they are beyond the scope of our work and are left for future study.

We focus our model development on representing two types of climate policies that can be implemented to decarbonize energy systems. One is an explicit cap on carbon emissions while the other is a carbon price. In this way our model can represent capacity expansion under current Eu-

ropean climate policy and policies that have been contemplated in the United States. Our model is also designed to accommodate a variety of technical approaches to decarbonization. This includes the use of renewable technologies, such as wind, photovoltaic (PV) solar, and concentrating solar power (CSP), nuclear generation, and fossil-fueled generation with carbon capture and sequestration (CCS) systems. Our model allows both AC and High-Voltage DC (HVDC) transmission lines, as well as stand-alone energy storage systems and integrating thermal energy storage (TES) into CSP plants. We demonstrate our model using a case study that is based on the Texas electric power system, using 2050 as the target planning year. Given the long planning horizon, we can treat the system as having no generation capacity (insomuch as existing capacity today will likely be retired by 2050) but having existing transmission capacity (as such assets are long-lived). By changing the stringency of the decarbonization targets, we examine how the generation and transmission mixes change. We also examine a number of sensitivity cases, which show the impacts of changes in technology development and transmission congestion on the decarbonization pathway of the system. Although our results are specific to our Texas-based case study, we make a number of qualitative observations that are likely broadly applicable. Moreover, our modeling framework is sufficiently generic that it can be applied to determining a decarbonization pathway for any power system.

Our work makes two primary contributions to the existing literature. First, we propose a stochastic optimization framework that represents multi-scale uncertainties, AC and HVDC transmission, renewable generators, and energy-storage facilities. We also incorporate a novel representation of the operation of TES that is incorporated into CSP plants in our model. Second, through our case study and sensitivity analyses, we examine the impacts of a variety of policy and technical factors on power system decarbonization. Specifically, we study the effects of environmental policies, investment mix, technology improvements, demand growth, and transmission-network congestion. We also examine tradeoffs between the use of wind and solar generation for decarbonization.

The remainder of this chapter is organized as follows. Section 2.2 details the proposed planning model. Section 2.3 summarizes our case-study data, providing more details on the technology options that are modeled in our analysis. Section 2.4 summarizes our case-study results in a base case, while section 2.5 provides the results of our sensitivity analyses. Section 2.6 concludes.

2.2 Model

Our model follows the approach that is used by Liu et al. (2018b) in representing capacity-planning decisions. More specifically, we use a static two-stage stochastic investment model. The first stage concerns all of the generation- and transmission-investment decisions. The second stage captures all of the operating conditions. Long-term/large-scale uncertainties are represented explicitly in scenarios that define the second stage while short-term/small-scale uncertainties are represented implicitly via different representative operating conditions.

We proceed in this section by first providing model notation, which is then followed by a detailed model formulation. We then further describe the technical and economic features that are captured in the model.

2.2.1 Model Notation

Indices

c	Index of thermal generation units.
g	Index of wind generation units.
v	Index of PV generation units.
k	Index of CSP generation units.
r	Index of BES units.
s	Index of stand-alone energy-storage units.
l	Index of existing and candidate transmission lines.
o	Index of operating conditions.
n	Index of power system nodes.
ref	Index for reference node.
t	Index of time periods within an operating condition.
w	Index of scenarios.
τ	Index of final time period within an operating condition.

Sets

$\zeta^s(l)$	Sending-end node of transmission line l .
$\zeta^r(l)$	Receiving-end node of transmission line l .
Ω_n^C	Set of thermal units that are located at node n .
Ω_n^G	Set of wind units that are located at node n .
Ω_n^V	Set of PV units that are located at node n .
Ω_n^K	Set of CSP units that are located at node n .
Ω^{L+}	Set of candidate transmission lines.
Ω_a^{L+}	Set of candidate AC transmission lines.
Ω_n^R	Set of BES units that are located at node n .
Ω_n^S	Set of energy-storage units that are located at node n .

We divide the transmission lines into existing and candidate lines. This is because of our assumption that transmission-infrastructure investments are sufficiently long-lived that transmission corridors that are existing today are still operational in the future target year of the analysis (*i.e.*, 2050 in our case study). We further subdivide the transmission lines into two technology types– AC and HVDC. The reason for this division is that power flows along the two types of lines are modeled differently. We model power flows along AC transmission lines using a linear approximation of Kirchhoff’s laws. Power flows along HVDC lines are represented using a ‘pipeline assumption,’ whereby power flows can be directed along each line. This is because each HVDC line is a radial connection between a single point in the AC network to a candidate wind or CSP site.

We assume that there are no existing generation or energy-storage units in the target year of the analysis. This is because such assets are too short-lived for units that are operational today to be available in the target year. Thus, all thermal, wind, PV, CSP, and energy-storage units are in actuality candidate units that must be built to be available for operational use. As we further detail below, we model three types of energy-storage units. One are TES systems, which are directly integrated into CSP plants and can only be used to store thermal energy that is produced by the solar field of the respective CSP. The second are BES units, which are assumed to be built in the same locations as the candidate PV plants. The other are stand-alone energy-storage units, which can be used to store electrical energy from any source. The index, s , corresponds to such stand-alone energy-storage units.

Parameters

$a_{v,o,t}$	Hour- t capacity factor of PV unit v in operating condition o [p.u.].
$b_{k,o,t}$	Hour- t capacity factor of CSP unit k in operating condition o [p.u.].
B_l	Susceptance of transmission line l [S].
$\bar{c}_k$	Energy-storage capacity of TES system of CSP unit k [h].
$\bar{c}_s$	Energy-storage capacity of energy-storage unit s [h].
$\bar{c}_r$	Energy-storage capacity of BES unit r [h].
$f_{g,o,t}$	Hour- t capacity factor of wind unit g in operating condition o [p.u.].
$F_l^{\max}$	Capacity of transmission line l [MW].
I_c^C	Investment cost of thermal unit c [\$/MW].
$I^{C,\max}$	Investment budget for building thermal units [\$].
I_g^G	Investment cost of wind unit g [\$/MW].
$I^{G,\max}$	Investment budget for building wind units [\$].
I_v^V	Investment cost of PV unit v [\$/MW].
$I^{V,\max}$	Investment budget for building PV units [\$].
I_k^K	Investment cost of CSP unit k [\$/MW].
$I^{K,\max}$	Investment budget for building CSP units [\$].
I_r^R	Investment cost of BES unit r [\$/MW].
$I^{R,\max}$	Investment budget for building BES units [\$].
I_s^S	Investment cost of energy-storage unit s [\$/MW].
$I^{S,\max}$	Investment budget for building energy-storage units [\$].
I_l^L	Investment cost of candidate transmission line l [\$].
$I^{L,\max}$	Investment budget for building candidate transmission lines [\$].
M	A large fixed constant.
$P_{n,o,w,t}^{D,\max}$	Hour- t load at node n in operating condition o of scenario w [MW].
$\bar{P}_c^{C,\max}$	Maximum capacity of thermal unit c that can be built [MW].
$\bar{P}_g^{G,\max}$	Maximum capacity of wind unit g that can be built [MW].
$\bar{P}_k^{K,\max}$	Maximum capacity of CSP unit k that can be built [MW].
$\bar{P}_r^{R,\max}$	Maximum capacity of BES unit r that can be built [MW].
$\bar{P}_s^{S,\max}$	Maximum capacity of energy-storage unit s that can be built [MW].
$\bar{P}_v^{V,\max}$	Maximum capacity of PV unit v that can be built [MW].
β_k	Hourly energy-retention rate of TES system in CSP unit k [p.u.].
β_r	Hourly energy-retention rate of BES unit r [p.u.].
$\bar{\epsilon}$	Carbon-emissions limit [t].

ϵ_c	Carbon-emissions rate of thermal unit c [t/MWh].
ϵ_s	Carbon-emissions rate of energy-storage unit c [t/MWh].
η_k^K	Charging efficiency of TES system in CSP unit k [p.u.].
η_r^R	Roundtrip efficiency of BES unit r [p.u.].
η_s^S	Roundtrip efficiency of energy-storage unit s [p.u.].
κ_c^C	Production cost of thermal unit c [\$/MWh].
κ^D	Load-shedding cost [\$/MWh].
κ_g^G	Production cost of wind unit g [\$/MWh].
κ_k^K	Production cost of CSP unit k [\$/MWh].
$\kappa_k^{K,P}$	Charging cost of TES system of CSP unit k [\$/MWh].
$\kappa_k^{K,T}$	Discharging cost of TES system of CSP unit k [\$/MWh].
$\kappa_r^{R,P}$	Charging cost of BES unit r [\$/MWh].
$\kappa_r^{R,T}$	Discharging cost of BES unit r [\$/MWh].
$\kappa_s^{S,P}$	Charging cost of energy-storage unit s [\$/MWh].
$\kappa_s^{S,T}$	Discharging cost of energy-storage unit s [\$/MWh].
κ_v^V	Production cost of PV unit v [\$/MWh].
ξ_c^d	Ramp-down limit of thermal unit c [MW/h].
ξ_c^u	Ramp-up limit of thermal unit c [MW/h].
ρ_o	Weight of operating condition o [days].
ϕ_w	Probability of scenario w .
χ	Carbon-emissions price [\$/t].

Decision Variables

$p_{c,o,w,t}^C$	Hour- t production level of thermal unit c in operating condition o of scenario w [MW].
$p_c^{C,\max}$	Capacity of thermal unit c that is built [MW].
$p_{n,o,w,t}^D$	Hour- t load at node n that is shed in operating condition o of scenario w [MW].
$p_{g,o,w,t}^G$	Hour- t production level of wind unit g in operating condition o of scenario w [MW].
$p_g^{G,\max}$	Capacity of wind unit g that is built [MW].
$p_{v,o,w,t}^V$	Hour- t production level of PV unit v in operating condition o of scenario w [MW].
$p_v^{V,\max}$	Capacity of PV unit v that is built [MW].

$p_{k,o,w,t}^K$	Hour- t production level of CSP unit k in operating condition o of scenario w [MW].
$p_k^{K,\max}$	Capacity of CSP unit k that is built [MW].
$p_r^{R,\max}$	Capacity of BES unit r that is built [MW].
$p_s^{S,\max}$	Capacity of energy-storage unit s that is built [MW].
$p_{l,o,w,t}^L$	Hour- t power flow through transmission line l in operating condition o of scenario w [MW].
x_l^L	Binary variable that equals 1 if candidate transmission line l is built and equals 0 otherwise.
$\gamma_{k,o,w,t}^{K,L}$	Ending hour- t state of charge (SOC) of TES system in CSP unit k in operating condition o of scenario w [MWh]
$\gamma_{k,o,w,t}^{K,P}$	Hour- t charging rate of TES system in CSP unit k in operating condition o of scenario w [MW].
$\gamma_{k,o,w,t}^{K,T}$	Hour- t discharging rate of TES system in CSP unit k in operating condition o of scenario w [MW].
$\gamma_{r,o,w,t}^{R,L}$	Ending hour- t SOC of BES unit r in operating condition o of scenario w [MWh].
$\gamma_{r,o,w,t}^{R,P}$	Hour- t charging rate of BES unit r in operating condition o of scenario w [MW].
$\gamma_{r,o,w,t}^{R,T}$	Hour- t discharging rate of BES unit r in operating condition o of scenario w [MW].
$\gamma_{s,o,w,t}^{S,L}$	Ending hour- t SOC of energy-storage unit s in operating condition o of scenario w [MWh].
$\gamma_{s,o,w,t}^{S,P}$	Hour- t charging rate of energy-storage unit s in operating condition o of scenario w [MW].
$\gamma_{s,o,w,t}^{S,T}$	Hour- t discharging rate of energy-storage unit s in operating condition o of scenario w [MW].
$\theta_{n,o,w,t}$	Hour- t phase angle at node n in operating condition o of scenario w [rad].

2.2.2 Model Formulation

Our proposed two-stage stochastic planning model is formulated as:

$$\begin{aligned}
 \min \quad & \sum_c I_c^C p_c^{C,\max} + \sum_g I_g^G p_g^{G,\max} + \sum_v I_v^V p_v^{V,\max} + \sum_k I_k^K p_k^{K,\max} + \sum_r I_r^R p_r^{R,\max} + \sum_s I_s^S p_s^{S,\max} \\
 & + \sum_{l \in \Omega^L} I_l^L x_l^L + \sum_{w,o,t} \phi_w \rho_o \left\{ \sum_c (\kappa_c^C + \chi \epsilon_c) p_{c,o,w,t}^C + \sum_n \kappa_n^D p_{n,o,w,t}^D + \sum_g \kappa_g^G p_{g,o,w,t}^G \right. \\
 & \left. + \sum_k \kappa_k^K p_{k,o,w,t}^K + \sum_r \kappa_r^R p_{r,o,w,t}^R + \sum_s \kappa_s^S p_{s,o,w,t}^S \right\}
 \end{aligned} \tag{2.1}$$

$$\begin{aligned}
 & + \sum_k \left(\kappa_k^K p_{k,o,w,t}^K + \kappa_k^{K,P} \gamma_{k,o,w,t}^{K,P} + \kappa_k^{K,T} \gamma_{k,o,w,t}^{K,T} \right) + \sum_r \left[\kappa_r^{R,P} \gamma_{r,o,w,t}^{R,P} + \kappa_r^{R,T} \gamma_{r,o,w,t}^{R,T} \right] \\
 & + \sum_s \left[\kappa_s^{S,P} \gamma_{s,o,w,t}^{S,P} + \left(\kappa_s^{S,T} + \chi \epsilon_s \right) \gamma_{s,o,w,t}^{S,T} / \eta_s^S \right] + \sum_v \kappa_v^V p_{v,o,w,t}^V \Big\} \quad (2.2)
 \end{aligned}$$

$$\text{s.t. } 0 \leq p_c^{C,\max} \leq \bar{p}_c^{C,\max}, \quad \forall c \quad (2.3)$$

$$0 \leq p_g^{G,\max} \leq \bar{p}_g^{G,\max}, \quad \forall g \quad (2.4)$$

$$0 \leq p_v^{V,\max} \leq \bar{p}_v^{V,\max}, \quad \forall v \quad (2.5)$$

$$0 \leq p_k^{K,\max} \leq \bar{p}_k^{K,\max}, \quad \forall k \quad (2.6)$$

$$0 \leq p_r^{R,\max} \leq \bar{p}_r^{R,\max}, \quad \forall r \quad (2.7)$$

$$0 \leq p_s^{S,\max} \leq \bar{p}_s^{S,\max}, \quad \forall s \quad (2.8)$$

$$x_l^L \in \{0, 1\}, \quad \forall l \in \Omega^{L+} \quad (2.9)$$

$$\sum_c I_c^C p_c^{C,\max} \leq I^{C,\max} \quad (2.10)$$

$$\sum_g I_g^G p_g^{G,\max} \leq I^{G,\max} \quad (2.11)$$

$$\sum_v I_v^V p_v^{V,\max} \leq I^{V,\max} \quad (2.12)$$

$$\sum_k I_k^K p_k^{K,\max} \leq I^{K,\max} \quad (2.13)$$

$$\sum_r I_r^R p_r^{R,\max} \leq I^{R,\max} \quad (2.14)$$

$$\sum_s I_s^S p_s^{S,\max} \leq I^{S,\max} \quad (2.15)$$

$$\sum_{l \in \Omega^{L+}} I_l^L x_l^L \leq I^{L,\max} \quad (2.16)$$

$$\sum_{w,o,t} \phi_w \rho_o \left(\sum_c \epsilon_c p_{c,o,w,t}^C + \sum_s \epsilon_s \gamma_{s,o,w,t}^{S,T} \right) \leq \bar{\epsilon} \quad (2.17)$$

$$\begin{aligned}
 & \sum_{c \in \Omega_n^C} p_{c,o,w,t}^C + \sum_{g \in \Omega_n^G} p_{g,o,w,t}^G + \sum_{k \in \Omega_n^K} \left(p_{k,o,w,t}^K - \gamma_{k,o,w,t}^{K,P} + \gamma_{k,o,w,t}^{K,T} \right) + \sum_{r \in \Omega_n^R} \left(-\gamma_{r,o,w,t}^{R,P} + \gamma_{r,o,w,t}^{R,T} \right) \\
 & - \sum_{s \in \Omega_n^S} \left(\gamma_{s,o,w,t}^{S,P} - \gamma_{s,o,w,t}^{S,T} / \eta_s^S \right) + \sum_{v \in \Omega_n^V} p_{v,o,w,t}^V + \sum_{l, \zeta^r(l)=n} p_{l,o,w,t}^L - \sum_{l, \zeta^s(l)=n} p_{l,o,w,t}^L = P_{n,o,w,t}^{D,\max} - p_{n,o,w,t}^D,
 \end{aligned} \quad (2.18)$$

$$\forall n, o, w, t$$

$$0 \leq p_{c,o,w,t}^C \leq p_c^{C,\max}, \quad \forall c, o, w, t \quad (2.19)$$

$$-\xi_c^d \leq p_{c,o,w,t}^C - p_{c,o,w,t-1}^C \leq \xi_c^u, \quad \forall c, o, w, t > 1 \quad (2.20)$$

$$0 \leq p_{g,o,w,t}^G \leq f_{g,o,t} p_g^{G,\max}, \quad \forall g, o, w, t \quad (2.21)$$

$$0 \leq p_{v,o,w,t}^V \leq a_{v,o,t} p_v^{V,\max}, \quad \forall v, o, w, t \quad (2.22)$$

$$0 \leq p_{k,o,w,t}^K + \gamma_{k,o,w,t}^{K,P} \leq b_{k,o,t} p_k^{K,\max}, \quad \forall k, o, w, t \quad (2.23)$$

$$0 \leq \gamma_{k,o,w,t}^{K,P} \leq b_{k,o,t} p_k^{K,\max}, \quad \forall k, o, w, t \quad (2.24)$$

$$0 \leq p_{k,o,w,t}^K + \gamma_{k,o,w,t}^{K,T} \leq p_k^{K,\max}, \quad \forall k, o, w, t \quad (2.25)$$

$$0 \leq \gamma_{k,o,w,t}^{K,T} \leq p_k^{K,\max}, \quad \forall k, o, w, t \quad (2.26)$$

$$\gamma_{k,o,w,t}^{K,L} = \beta_k \gamma_{k,o,w,t-1}^{K,L} + \eta_k^K \gamma_{k,o,w,t}^{K,P} - \gamma_{k,o,w,t}^{K,T}, \quad \forall k, o, w, t \quad (2.27)$$

$$0 \leq \gamma_{k,o,w,t}^{K,L} \leq \bar{c}_k p_k^{K,\max}, \quad \forall k, o, w, t \quad (2.28)$$

$$0 \leq \gamma_{r,o,w,t}^{R,P} \leq p_r^{R,\max}, \quad \forall r, o, w, t \quad (2.29)$$

$$0 \leq \gamma_{r,o,w,t}^{R,T} \leq p_r^{R,\max}, \quad \forall r, o, w, t \quad (2.30)$$

$$\gamma_{r,o,w,t}^{R,L} = \beta_r \gamma_{r,o,w,t-1}^{R,L} + \eta_r^R \gamma_{r,o,w,t}^{R,P} - \gamma_{r,o,w,t}^{R,T}, \quad \forall r, o, w, t \quad (2.31)$$

$$0 \leq \gamma_{r,o,w,t}^{R,L} \leq \bar{c}_r p_r^{R,\max}, \quad \forall r, o, w, t \quad (2.32)$$

$$\gamma_{r,o,w,\tau}^{R,L} \geq \gamma_{r,o,w,0}^{R,L} = 0.5 p_r^{R,\max}, \quad \forall r, o, w \quad (2.33)$$

$$0 \leq \gamma_{s,o,w,t}^{S,P} \leq p_s^{S,\max}, \quad \forall s, o, w, t \quad (2.34)$$

$$0 \leq \gamma_{s,o,w,t}^{S,T} / \eta_s^S \leq p_s^{S,\max}, \quad \forall s, o, w, t \quad (2.35)$$

$$\gamma_{s,o,w,t}^{S,L} = \gamma_{s,o,w,t-1}^{S,L} + \gamma_{s,o,w,t}^{S,P} - \gamma_{s,o,w,t}^{S,T}, \quad \forall s, o, w, t \quad (2.36)$$

$$0 \leq \gamma_{s,o,w,t}^{S,L} \leq \bar{c}_s p_s^{S,\max}, \quad \forall s, o, w, t \quad (2.37)$$

$$\gamma_{s,o,w,\tau}^{S,L} \geq \gamma_{s,o,w,0}^{S,L} = 0.5 p_s^{S,\max}, \quad \forall s, o, w \quad (2.38)$$

$$-F_l^{\max} \leq p_{l,o,w,t}^L \leq F_l^{\max}, \quad \forall l \notin \Omega^{L+}, o, w, t \quad (2.39)$$

$$-F_l^{\max} x_l^L \leq p_{l,o,w,t}^L \leq F_l^{\max} x_l^L, \quad \forall l \in \Omega^{L+}, o, w, t \quad (2.40)$$

$$p_{l,o,w,t}^L = B_l \cdot (\theta_{\zeta^s(l),o,w,t} - \theta_{\zeta^r(l),o,w,t}), \quad \forall l \notin \Omega^{L+}, o, w, t \quad (2.41)$$

$$-(1 - x_l^L) M \leq p_{l,o,w,t}^L - B_l \cdot (\theta_{\zeta^s(l),o,w,t} - \theta_{\zeta^r(l),o,w,t}) \leq (1 - x_l^L) M, \quad \forall l \in \Omega_a^{L+}, o, w, t \quad (2.42)$$

$$-\pi \leq \theta_{n,o,w,t} \leq \pi, \quad \forall n, o, w, t \quad (2.43)$$

$$\theta_{\text{ref},o,w,t} = 0, \quad \forall o, w, t \quad (2.44)$$

$$0 \leq p_{n,o,w,t}^D \leq P_{n,o,w,t}^{D,\max}, \quad \forall n, o, w, t. \quad (2.45)$$

Objective function (2.1) consists of total investment and expected operation cost of the power system. We assume that the investment-cost parameters, I_c^C , I_g^G , I_r^R , I_k^K , I_v^V , I_s^S , I_l^L , are annualized to make the investment and operations costs comparable to one another. Otherwise, the costs would be skewed because the investments are long-lived while operations may only be modeled over a short duration (*e.g.*, for a single representative year in our case study). The simplest way to annualize the investment costs is to multiply the overnight cost of building each technology by:

$$\frac{i \cdot (1+i)^y}{(1+i)^y - 1},$$

where y is the assumed lifetime of the asset and i is the real interest rate.

Operating costs are computed under each hour of each operating condition of each scenario. These operating costs include the costs of producing energy from the thermal, wind, PV, and CSP units, the cost of operating the BES, the TES that is embedded in the CSP units and the stand-alone energy-storage units, and the cost of any load curtailment.

We also include operating costs that are associated with CO₂ emissions (*e.g.*, due to a Pigouvian-tax or cap-and-trade policy). We assume two possible sources of CO₂ emissions—thermal generators and stand-alone energy-storage units. Any fossil-fueled thermal generator that does not have a CCS system incorporated in it will release CO₂. The stand-alone energy-storage technology that we focus on in our case study is diabatic compressed-air energy storage (CAES). As discussed by Succar et al. (2006); Greenblatt et al. (2007); Succar and Williams (2008), diabatic CAES is in actuality a hybrid energy-storage system that combusts natural gas during the discharging cycle. Thus, we allow for cases in which the use of the energy-storage technology can result in CO₂ emissions. We also examine, in one of our sensitivity cases in section 2.5.5, the viability of adiabatic CAES, which is a pure energy-storage technology, the use of which does not entail any direct CO₂ emissions.

The model has two types of constraints. Constraints (2.3)–(2.17) pertain to the first (investment) stage while the remaining pertain to the second (operating) stage. Constraints (2.3)–(2.8) impose capacity limits on how much thermal, wind, PV, CSP, BES, and energy-storage capacity can be built at each candidate location. These constraints may reflect land-use, resource, or other physical limitations on the deployment of these units. Constraints (2.9) impose the binary nature of transmission-expansion decisions, which have this ‘lumpy’ nature. Constraints (2.10)–(2.16) impose budget constraints on the investment decisions. Constraint (2.17) imposes an expected carbon-emissions limit (*e.g.*, due to a cap-and-trade system limiting total system emissions). As such, our model captures two common means of imposing carbon-emissions limits. One is an explicit limit, as imposed by

constraint (2.17), while the other is a more implicit cost or tax on carbon emissions, which is captured in objective function (2.1). In practice, policy makers often focus on implementing one of these types of policy mechanisms as opposed to both.

Constraints (2.18) impose nodal load balance in each operating period. Constraints (2.19) and (2.20) impose capacity and ramping limits, respectively, on the operation of the thermal units. Constraints (2.21) and (2.22) impose capacity constraints on the output of the wind and PV units, respectively. The energy that is available to be produced in each operating period is represented using a capacity factor, which captures the impact of weather conditions. The same approach is used to represent the impact of weather on energy that is available from CSP plants.

Constraints (2.23)–(2.28) pertain to the operation of the CSP units. We model the CSP units by adapting the methodology that is developed by Sioshansi and Denholm (2010a,b). Constraints (2.23) restrict the total amount of energy that each CSP unit produces and stores in its TES system to be no greater than the amount of energy that is collected by the unit's solar field. The energy that is collected by a CSP plant's solar field is thermal. As such, we apply the assumed powerblock efficiency to convert the thermal energy into the equivalent amount of electrical energy that would be produced by the CSP unit when the stored energy is discharged. Doing so allows for more streamlined modeling of the CSP unit and its integrated TES system within a planning model. Constraints (2.24) ensure that the TES system does not store more than its power capacity allows. Constraints (2.25) limit the total production of each CSP unit, either by using energy directly from its solar field or from its TES system, to be no greater than the power capacity of its powerblock. Constraints (2.26) limit the amount of energy that is discharged from the TES system based on its power capacity. Constraints (2.27) define how the SOC of the TES system in each CSP unit evolves from one operating period to the next. Finally, constraints (2.28) impose energy limits on the SOC of the TES system in each CSP unit.

Constraints (2.29)–(2.33) pertain to the operation of the BES units. Constraints (2.29)–(2.30) set power-capacity restrictions on the charging and discharging, respectively, of BES units. Constraints (2.31) define how the SOC of each BES unit varies from one operating period to the next, while constraints (2.32) enforce energy limits on the SOC. Constraints (2.33) fix the SOC of each BES unit at the beginning of each operating condition (which are taken to be representative operating days in our case study) to 50% of its energy capacity. It is further imposed that the SOC at the end of each operating condition to be above this level.

Constraints (2.34)–(2.38) describe the operation the stand-alone energy-storage units, whose modeling is similar to that adopted for BES installations. Constraints (2.34)–(2.35) impose power-

capacity limits on the charging and discharging, respectively, of energy-storage units. Constraints (2.36) define how the SOC of each energy-storage unit changes from one operating period to the next and constraints (2.37) impose energy limits on the SOC. Constraints (2.38) fix the SOC of each energy-storage unit at the beginning of each operating condition (which are taken to be representative operating days in our case study) to 50% of its energy capacity. It is further enforced that the SOC at the end of each operating condition to be above this level. Graves et al. (1999) suggest this as a heuristic technique to ensure that a model does not fully deplete energy storage at the end of each operating horizon. The modeling of the stand-alone energy-storage units is based on the work of Drury et al. (2011), who develop an optimization model that is tailored to CAES.

Constraints (2.39) impose power-flow limits on each existing transmission line while constraints (2.40) do the same for candidate lines. These latter constraints also force the power flow on a candidate transmission line that is not built to equal zero, which is a physical requirement. Constraints (2.41) define power flows on existing transmission lines in terms of differences in the phase angles at the two ends of the line. Constraints (2.42) define power flows on candidate transmission lines analogously. These latter constraints are written in this inequality form to allow the flows on a candidate transmission line that is not built to equal zero without forcing the phase angles at the two ends of that line to equal one another. This is a standard approach to modeling power flows with binary transmission-investment decisions. Conejo et al. (2016) provide further details on how these constraints are modeled. Power flows in HVDC lines are not modeled using phase angles. Rather, because these lines are radial connections to distant candidate wind or CSP sites, their power flows can be set arbitrarily within their bounds without impacting power flows elsewhere in the network. Constraints (2.43) impose limits on the phase angles and constraints (2.44) fix the phase angles at the reference node to be zero.

Finally, constraints (2.45) limit the amount of load that is curtailed at each node to be no greater than the nodal demand.

Our proposed capacity-expansion model is a mixed-integer linear optimization problem. Thus, it can be tractably solved using off-the-shelf commercial software tools, such as CPLEX or Gurobi. Although we have a stochastic optimization problem, explicit non-anticipativity constraints are not needed. This is because the investment decisions (*i.e.*, $p_c^{C,\max}$, $p_g^{G,\max}$, $p_v^{V,\max}$, $p_k^{K,\max}$, $p_r^{R,\max}$, $p_s^{S,\max}$, and x_l^L) are defined so as not to depend on the realization of the second-stage scenarios, w . The proposed model can also be used in a rolling-horizon fashion over time, whereby future scenarios are periodically updated (*e.g.*, the model can be re-run every year with updated scenarios of future conditions). Domínguez et al. (2015) argue that using a two-stage stochastic planning

model in such a manner can yield solutions that are similar to those that are achieved using a more computationally costly multi-period model.

2.3 Case-Study Data

Our case study considers the 28-node system, which is a modified version of the IEEE 24-node test system, that is shown in Figure 2.1. Generating and energy-storage units are labeled in the figure based on how they are indexed in the model formulation (*i.e.*, c , g , v , k , r and s for thermal, wind, PV, CSP, BES, and stand-alone energy-storage units, respectively). Existing transmission lines are indicated by solid lines whereas candidate AC lines are indicated by dashed lines. Dotted lines indicate candidate HVDC lines. Loads are indicated by arrows. The topology of the network and other system data are adapted from the Electric Reliability Council of Texas (ERCOT) power system. However, loads, generator, and transmission-lines capacities are scaled down so as to keep the references loads at 3.7 GW in the base case. We focus our case study on ERCOT because it is a mostly electrically isolated power system (with the exception of some limited DC tie lines to neighboring interconnects) with substantial load and potential for the deployment of renewables. Twenty-four of the nodes represent the center of the network where the loads and most of the candidate generating and energy-storage units are located. As such, these nodes are connected by existing AC lines and candidate AC lines can be built to reinforce connections within the center of the network. The remaining nodes are isolated and can be connected to the center of the network with candidate HVDC lines. Two of the nodes are candidate locations for offshore wind generators while the other two are candidate locations for CSP plants.

Our case study includes 83 candidate units: 40 thermal, 11 onshore wind, ten PV, two offshore wind, two CSP, ten BES, and eight stand-alone energy-storage units. Additional candidate CSP units are considered in one of the sensitivity cases that is examined in section 2.5.3. We consider ten different thermal generation technologies: advanced pulverized coal units (with and without CCS), conventional natural gas-fired combined cycle (NGCC) units, advanced NGCC units (with and without CCS), conventional and advanced natural gas-fired combustion turbines (CT), integrated gasification combined cycle (IGCC) units (with and without CCS), and advanced nuclear units. This represents generating technologies that, as reported by the United States Energy Information Administration (EIA),¹ are available today or that are anticipated to be commercially

¹<https://www.eia.gov/analysis/studies/powerplants/capitalcost/>

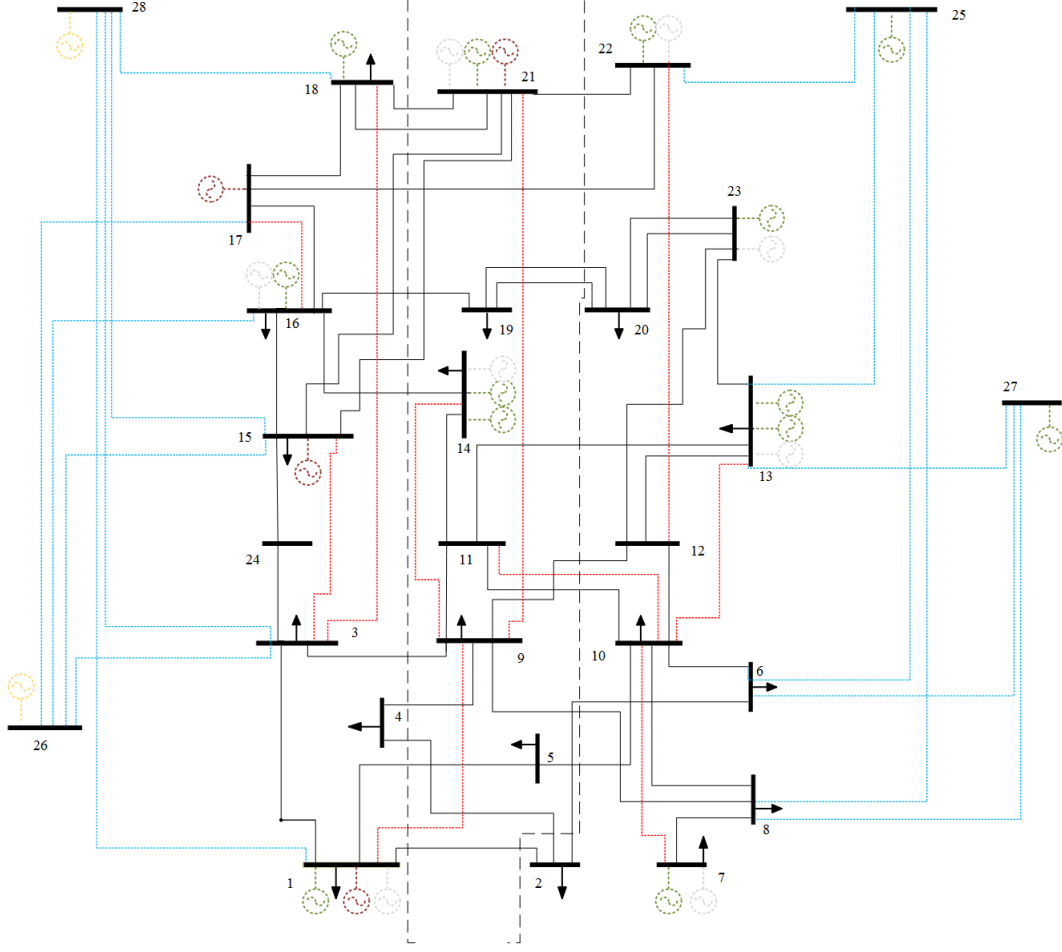


Figure 2.1: 28-Node Network Diagram

available by the target year of our analysis (2050).

Table 2.1 summarizes the cost, maximum-capacity, and emissions-rate data for the thermal units, which are obtained from the EIA. All of the technologies are assumed to be able to ramp over their full output range, with the exception of nuclear units, which are assumed to have no ramping capability. Instead, nuclear units are assumed to continuously operate at their installed capacity as must-run units. The operating costs that are reported in Table 2.1 are computed using projections of future fuel prices that are published by U.S. Energy Information Administration (2017). The maximum capacity for each technology that is reported in Table 2.1 is the total amount that can be installed across all of the candidate locations where that technology can be built.

There are 11 candidate locations, which are in the center of the test system, for onshore wind.

Table 2.1: Data For Candidate Thermal Units

Technology	Investment Cost [\$ /kW]	Operating Cost [\$ /MWh]	Maximum Capacity [MW]	ϵ_c
Advanced Pulverized Coal Without CCS	2933	25.95	3800	0.841
Advanced Pulverized Coal With CCS	5227	38.15	3800	0.112
Conventional NGCC	917	46.82	2880	0.374
Advanced NGCC Without CCS	1023	42.69	2400	0.341
Advanced NGCC With CCS	2095	52.91	2160	0.041
Conventional CT	973	81.97	1600	0.576
Advanced CT	676	70.14	2400	0.518
IGCC Without CCS	3784	26.19	3800	0.743
IGCC With CCS	6599	33.22	3000	0.097
Advanced Nuclear	5530	16.06	4800	0.000

Conversely, offshore wind can only be built at two nodes that are isolated from the center of the system. Onshore wind has investment costs that range between \$1970/kW and \$2100/kW and operating costs that range between \$2.00/MWh and \$2.71/MWh. A maximum of 7700 MW of wind can be built at the onshore sites. Up to 4000 MW of offshore wind can be built. These units have higher investment costs that range between \$2900/kW and \$3190/kW, which reflects the greater complexity that installing offshore wind entails. These units also have higher operating costs that range between \$4.85/MWh and \$5.35/MWh, which reflects the greater challenges in their operation and maintenance. One of the candidate offshore wind units is closer to the shore, which results in its having a lower investment cost (and a lower cost for building the associated HVDC line that connects it to the center of the network). The other offshore wind unit is further from the shore, giving it a higher average capacity factor and lower operating cost. Investment and operating costs of the wind units are obtained from the EIA and the work of Krohn et al. (2009).

A maximum of 7000 MW of PV can be built in the center of the test system, including 2100 MW of rooftop installations. Investment costs of the PV units for rooftop installations range between \$1000/kW and \$1012/kW, while those of the PV plants for utility scale are between \$790/kW and \$801/kW. Operating costs range between \$1.50/MWh and \$1.52/MWh. DC to AC ratio is assumed equal to 1.2 for each unit while inverter efficiency is set to 96%.

Up to 2000 MW total of CSP capacity can be built across the two candidate locations. We assume that the solar field of each CSP plant is scaled to maintain a solar multiple of 2.0. A CSP plant with a solar multiple of 2.0 is sized to provide sufficient thermal energy to operate the powerblock at its

rated capacity with a direct normal irradiance of 950 W/m^2 , a 5-m/s wind speed, and a 25-C ambient temperature. We, furthermore, assume that each CSP plant has a TES system with $\bar{c}_k = 4$ hours of storage capacity. This means that a fully charged TES system can be discharged and operated at its nameplate capacity for four consecutive hours.² Operating and investment costs for the CSP units are obtained from version 2017.1.17 of the System Advisor Model (SAM). Blair et al. (2014) provide a general overview and description of this software tool. We assume power-tower CSP systems with investment costs ranging between $\$3272/\text{kW}$ and $\$3599/\text{kW}$ and generation costs ranging between $\$3.80/\text{MWh}$ and $\$4.00/\text{MWh}$. The TES systems in the CSP plants are assumed to have operating costs of $\$0.1400/\text{MWh}$ in charging mode and $\$0.1549/\text{MWh}$ in discharging mode. We assume that the TES systems have a 99% roundtrip efficiency and 0.1% hourly thermal-energy losses, meaning that we have $\eta_k^K = 0.99$ and $\beta_k = 0.999$ for all k . As with the offshore wind units, one of the candidate CSP units is assumed to be closer to a load pocket in the system, meaning that this unit has a lower cost of being connected to the system with HVDC lines. However, this unit has a lower average capacity factor than the other CSP unit.

Each BES unit is assumed to have a fixed $\bar{c}_r = 4$ hours of storage capacity and up to 350 MW of power capacity. The related investment costs and physical properties are taken from the work of Fu et al. (2018). Investment costs range between $\$1960/\text{kW}$ and $\$2100/\text{kW}$ and are scaled to account for the actual usable SOC of the battery which ranges between 20% and 80% of the nominal battery capacity. Each unit has an operating cost of $\$0.50/\text{MWh}$ when charging energy and $\$1/\text{MWh}$ when discharging. Each BES unit is also assumed to have 99% roundtrip efficiency and 0.1% hourly energy loss, i.e. $\beta_r = 0.99$ and $\beta_r = 0.999$ for all r .

We model a single stand-alone energy-storage technology, which is diabatic CAES. Diabatic CAES is a hybrid energy-storage technology that uses both electricity (to compress air) and natural gas (when stored energy is discharged). When charged, electricity is used to drive a compressor, which injects pressurized air into a reservoir. When discharged, the compressed air is withdrawn from the reservoir and combined with natural gas in an open-cycle combustion turbine. Combustion produces high-pressure gas, which is expanded through a turbine, that drives both an electric generator and the air compressor. The benefit of using the compressed air in the discharging phase is that the combustion turbine has a heat rate of approximately 4200 BTU/kWh ,³ which is half that of

²In actuality, a fully charged TES system would operate very slightly below 100% of its nameplate capacity in the fourth hour. The reason for this is that the TES system incurs very slight thermal-energy losses from one hour to the next. Thus, the four-hour definition is based on an assumption of no thermal-energy losses.

³We report this heat rate in imperial (as opposed to SI) units, as BTU/kWh is still the most commonly used unit of generator efficiency in the United States.

a conventional CT. Because diabatic CAES is a hybrid energy-storage technology, only 0.7 MWh of stored electric energy must be used to discharge 1 MWh of electricity. Adiabatic CAES, further details of which are given in section 2.5.5, is considered in one of the sensitivity cases. Adiabatic CAES is a pure energy-storage technology, which does not entail the combustion of any fossil fuel when discharged.

Sioshansi et al. (2011) note that CAES is an attractive large-scale energy-storage technology with long life expectancy, low investment and maintenance costs, and reasonable roundtrip efficiency. Moreover, Succar and Williams (2008) find that many areas of the world that are potentially well suited for the deployment of CAES are also wind-rich. This includes ERCOT. This can be contrasted with the (currently) more commonly used pumped hydroelectric storage (PHS) technology, which requires access to water and reservoirs at different elevations. Water is scarce and the requisite geological formations are lacking in many wind-rich areas of the world, including in Texas.

We model eight candidate CAES units. Each CAES unit is assumed to have a fixed $\bar{c}_s = 10$ hours of storage capacity. Each CAES units can be built with up to 350 MW of power capacity. We take operating and investment costs and physical properties of the CAES plants from the work of Das et al. (2011). We assume that the investment costs range between \$550/kW and \$610/kW and that each CAES unit has an operating cost of \$0.50/MWh when charging energy. The discharging costs range between \$28.33/MWh and \$28.79/MWh, which primarily reflect the cost of natural gas. Each CAES unit is also assumed to have a carbon-emissions rate of 0.259 t/MWh and an efficiency of $\eta_s^S = 0.7$ (based on the characteristic that only 0.7 MWh of stored energy must be used to discharge 1 MWh of electricity).

The transmission system consists of 38 existing lines, which range in capacity from 200 MW to 800 MW and have an average capacity of about 440 MW. These existing lines are all AC and connect the nodes in the center of the network. Ten candidate AC lines, with capacities ranging between 400 MW and 800 MW, can be added to create further links amongst the nodes in the center of the network. There are a further 15 candidate HVDC lines, which only connect one of the four remote nodes to one of the nodes in the center of the network. Each of the distant nodes has multiple candidate HVDC lines with which it can be connected to the main grid, indicating that a particular offshore-wind or CSP project has multiple potential points of interconnection with the main network. Eight of the HVDC lines have a 2-GW capacity while the remainder have a 1-GW capacity. Transmission-investment costs are obtained from the work of Vaillancourt (2014).

We annualize the investment costs by assuming a 25-year depreciation period and a 10% real interest rate. This yields an 11% cost-annualization rate.

We model different load and wind- and solar-availability conditions using scenarios and operating conditions. Specifically, we represent operations in the target year through a set of operating conditions, which are indexed by o in the model. Each operating condition is taken to be a representative day, which is modeled at hourly time steps (*i.e.*, the index, t , corresponds to the 24 hours of each representative day). We use the scenarios, which are indexed by w in the model, to capture uncertainty in long-term demand growth between today and 2050.

To capture correlations in wind and solar availabilities and load, the operating-condition data are all obtained using historical data from the year 2012. Historical load data are obtained from ERCOT and scaled to obtain reference loads for the system. These reference loads are then apportioned to the nodes (*cf.* Figure 2.1) based on the historical distribution of loads to different zones of the ERCOT network.

Wind-speed data are obtained from the Wind Integration National Dataset (WIND) Toolkit. Draxl et al. (2015) provide details and meteorological validation of the WIND Toolkit. Wind speeds are used to determine the real-time availability of wind power, which is represented in the model through the capacity factors, $f_{g,o,t}$. King et al. (2014) validate the use of the WIND Toolkit for this type of simulation. We collect wind-speed data for 33 onshore locations and for the representative southwestern Texas offshore location. The wind-speed data are then processed using SAM, assuming the power curve of a 2-MW Vestas V80/2000 wind turbine with an 80-m hub height to generate hourly wind-availability data. Wind-availability profiles for the 13 candidate wind locations are obtained by spatial averaging of the profiles from the locations in the WIND Toolkit that are used.

Solar availabilities are simulated using SAM, based on weather data that are obtained from the National Solar Radiation Database (NSRDB). Sengupta et al. (2014) provide details and validation of the NSRDB data. We only consider two candidate CSP plants in our base case. As such, we use NSRDB data for five locations that are in relatively close proximity to one another to simulate the solar-availability of each CSP plant (*i.e.*, we use NSRDB data for 10 locations total). We use SAM to simulate the operation of a standard power-tower CSP plant with a 110-MW powerblock, a solar multiple of 2.0, and a TES system with 100 hours of storage capacity. From this simulation, we can determine the maximum amount of thermal energy that is captured by the CSP plant in each hour (*i.e.*, the 100-hour TES system ensures that no thermal energy is curtailed in the SAM simulation due to power constraints on the operation of the plant). We then use the assumed 41.2% average efficiency of the CSP plant's powerblock to convert the simulated thermal energy that is captured by the solar field in each hour to the amount of equivalent electrical energy that would be produced

by the powerblock. We finally normalize the potential electrical output by the 110-MW nameplate capacity of the powerblock to obtain hourly capacity factors for each CSP plant that is simulated in SAM. We again use spatial averaging to convert these capacity factors to the capacity factors for each of the two candidate CSP plants that are used in our planning model. Depending on ambient weather conditions, the capacity factor of a CSP plant can be greater than unity in some hours, indicating that the solar field collects more thermal energy than can be utilized by the powerblock. Such excess energy can be stored in the TES system. Solar availability and operation of PV units are simulated using SAM as described for CSP units.

Once we have a full set of load and wind- and solar-availability data for the year, we use k -means clustering to reduce the full year into a set of representative operating days. Thus, we obtain our final set of operating conditions (*i.e.*, the set over which the index, o , in our model is defined) from this process. MacQueen (1967) is a formative work that describes the k -means clustering algorithm, while Baringo and Conejo (2013b); Liu et al. (2018a) describe its use in obtaining operating conditions that represent the load, wind, and solar patterns of the year and respect their auto- and cross-correlations. The k -means clustering algorithm provides 15 clusters, each of which gives one representative day for the planning model (*i.e.*, our planning model has 15 day-long operating conditions). The load, wind, and solar data for the operating condition that is used to represent each cluster are obtained from the day within the cluster that is closest to the cluster centroid. We assign a weight to the representative day (*i.e.*, the value of ρ_o in the planning model) that is equal to the number of days within its associated cluster. This clustering results in the system having a peak load of 3.1 GW in the base scenario of the scenario tree.

In addition to the short-term uncertainty in load and renewably supply variability, we consider a source of long-term uncertainty, which is demand growth. Demand growth is represented explicitly via second-stage scenarios. We assume that with probability 0.5 the demands are equal to the values that are obtained from the k -means clustering ('no demand growth' scenario). With probability 0.3 demands are assumed to grow an additional 5% relative to this baseline value ('demand increase' scenario) and with probability 0.2 demands are 3% lower ('demand decrease' scenario). One reason for modeling uncertain demand growth is that it 'forces' the model to build sufficient capacity to serve higher-than-expected demands. This excess capacity that the demand-growth scenarios engender can be analogized to planning-reserve margins, which are normally enforced as constraints in deterministic capacity-expansion models. The uncertainty in demand growth also depends on the future integration of electric vehicles in the power system. For all these reasons, we also conduct a sensitivity analysis on the expected increase of the electricity consumption by

replacing the growth of 5% with growth values of 10%, 30%, and 50%, respectively (see section 2.5.6 for more details).

A benefit of modeling demand-growth scenarios (as opposed to planning-reserve margins) is that the objective function takes into account the expected cost of operating the system under higher-demand scenarios. Deterministic models that include planning-reserve constraints do not normally internalize the cost of operating the excess capacity that is built to provide the reserves (in the event that these units are needed to serve higher-than-expected demand). As such, the investments that are given by deterministic planning models with planning-reserve constraints tend to be skewed toward technologies that are overly costly to operate. Moreover, since the scope of the chapter is to provide a modeling framework for generation- and transmission-planning decisions in energy systems subject to decarbonization targets, we do not account for demand side management whose services can be detected in real-time markets and are similar to those provided by reserve margins. The online supplementary material includes tables that further detail our case-study data, which are excluded here for sake of brevity. All of the cases are programmed using version 21.1.2 of the GAMS mathematical programming language. They are solved using the hybrid branch-and-bound and cutting-plane algorithm in version 12.5.1.0 of the CPLEX solver on a system with Windows 10, a 2.6-GHz processor, and 8 GB of RAM.

2.4 Base-Case Results

In our analysis, we consider CSP, as opposed to PV, because CSP is a relatively dispatchable renewable resource owing to the low cost and high efficiency of integrating TES into a CSP plant. PV combined with BES units can provide similar services but at higher costs. Madaeni et al. (2012) show that this characteristic of CSP gives it considerable value for grid-integration purposes. Moreover, Liu et al. (2018b) conduct a long-term generation- and transmission-planning study of the ERCOT system, considering wind and PV solar as renewable-energy resources. They find that relatively little PV solar is built, by virtue of its relatively high energy cost, which is driven by its investment cost and low capacity factor. These properties of CSP *vis-à-vis* PV solar are consistent with the solar-insolation conditions of our case study. For all these reasons, we exclude the possibility to build PV and BES units in the base case, by setting the corresponding investment variables in model (2.1)-(2.45) equal to zero (see section 2.2.2). However, the trade-off related to the economical convenience of investing either in CPS with integrated TES or in PV combined with BES is explored in section 2.5.4.

Moreover, a system configuration without investments in PV and BES unit allows us to analyze situations similar to those foreseen by the Desertec project⁴ that aimed to integrate the electrical power systems of Europe and Northern Africa thanks to very expensive transmission lines passing through Italy and Spain. The project intended to develop investments in advanced renewable-based technologies in the Northern Africa, such as CPS plants, and to exchange this power with that produced by highly efficient onshore and offshore wind, respectively installed in the Northern Europe and in North Sea.

Figure 2.2 summarizes the mix of generation and energy-storage units that are built under a variety of cases with different carbon-emissions limits imposed on the system. The first case that is shown in the figure is a ‘business-as-usual’ (BAU) case, in which constraint (2.17) is relaxed and there is no carbon-emissions limit. The remaining cases impose constraint (2.17), with progressively more aggressive carbon-emissions limits that are defined as reductions relative to the BAU case. The final bar in figure 2.2 is labeled ‘100% Reduction’. With our case-study data and the technologies that we consider, complete decarbonization of the power system is extremely costly. As such, we allow minuscule carbon emissions in the base case and in all of the sensitivity cases that are considered in section 2.5. In all of the cases, carbon emissions are reduced by at least 99.7% relative to the BAU case. We, nevertheless, refer to these cases as having 100% carbon reductions. We do not impose a carbon tax in any of these cases, because we assume that explicit carbon-emissions limits are used as the policy mechanism to achieve decarbonization.

The BAU case yields investments solely in thermal and energy-storage units. This shows that absent any policy mechanism internalizing the social cost of carbon, thermal generation is the lowest-cost source of energy. Among the thermal-generation technologies that are available, natural gas-fired units are favored over other generating fuels. Most of the generating capacity in the BAU case is baseload NGCC, with some CAES and advanced CT capacity being built to serve on-peak loads.

As the carbon-emissions limit is enforced and made more stringent, the investment mix changes in a number of ways. 25% carbon-emissions reductions are achieved through greater investments in onshore wind. 50% and 75% carbon-emissions reductions are achieved using onshore wind and by shifting the thermal generation mix toward having CCS. As the thermal-generation mix shifts toward using CCS, there is still a preference for natural gas over coal as a generation fuel. These results show that onshore wind is initially the lowest-cost source of carbon abatement. CCS is only employed with more stringent carbon-reduction goals, once the prime onshore-wind resources are

⁴See <https://www.desertec.org/> for more information.

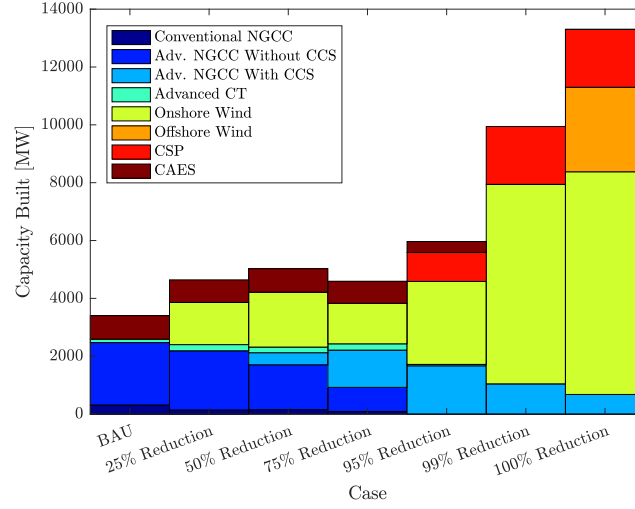


Figure 2.2: Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case

exploited. Notice that the capacity needed in the 75% reduction scenario of Figure 2.2 is lower than those constructed in the 50% and 25% cases, respectively. This result is driven by the trade-off between onshore wind units and NGCC plants with CCS. Notice that at 75% carbon emission reduction scenario, the CO₂ limit is strict but not so stringent to completely cut off conventional NGCC plants and NGCC units without CCS that are still built, even though in lower proportion compared to the 50% and the 25% reductions. On the other side, investments in NGCC plants with CCS significantly increase compared to 50% and the 25% reduction cases, which partially replace onshore wind installations. NGCC plants with CCS are more expensive with respect to wind units in terms of investment costs. However, they are more stable than wind units that, instead, have comparatively lower capacity factors. In this light, it is more convenient to use NGCC plants with CCS rather than building larger amount of wind plants. The final outcome is that the total amount of capacity built is lower than in the 50% and the 25% reduction cases because stable thermal units with CSS devices are preferred. This investment strategy does not abate system costs that, as expected, are higher than in the cases with lower emission reduction targets (see Figure 2.3). Greater carbon-emissions reductions beyond 75% require the use of CSP and offshore wind. This shows that the higher cost of these units, along with the cost of the HVDC connectors that are required to deliver their generation to load, makes these technologies quite expensive. As such, these units are only used with highly aggressive carbon-reduction targets. The fact that offshore wind and CSP are only built to achieve carbon-emissions reductions beyond 75% also implies that

HVDC lines are only built with these aggressive carbon-reduction goals.

Interestingly, we find that the role of CAES in providing system flexibility diminishes as the carbon-reduction target gets more aggressive. This is because of the carbon emissions that are associated with discharging CAES. As the carbon-emissions limit gets tighter, CAES is replaced with CSP (which has TES incorporated in it) as the primary source of supply-side flexibility. This diminishing use of CAES as the carbon-emissions limit becomes more stringent results in increasing wind-curtailment rates. Unlike onshore wind and CSP, offshore wind is only used in the most extreme case in which carbon emissions must be fully eliminated. The ‘low priority’ for building offshore wind stems from its relatively high cost, inflexibility (as TES cannot be incorporated in wind units), and it having little to no advantage over onshore wind in terms of real-time availability. Indeed, the only major benefit to offshore is that its availability is more stable. The distribution of hourly capacity factors of onshore wind has more extremes compared to offshore wind.

Figure 2.2 also shows that as the carbon-emissions limit gets more stringent, the total installed capacity increases considerably (*e.g.*, from 3.4 GW in the BAU case to 13 GW to achieve 100% carbon-emissions reductions). This is because of the lower capacity factor of wind and solar resources compared to dispatchable thermal generation. The 100%-carbon-reduction case results in a very small amount of NGCC with CCS being built. These units have lower emissions rates than CAES and operate only in a small handful of hours when CSP and wind plants are not fully able to serve the load. These small number of hours in which the NGCC units operate give the minuscule carbon emissions in the 100%-carbon-reduction case.

Although nuclear generation has no direct carbon emissions, nuclear capacity is not built in any of the carbon-reduction cases that we examine. This shows that nuclear is an uneconomic generation source in our case study, despite having no carbon footprint. One may attribute the lack of nuclear capacity to its extremely inflexible ramping capability that we assume. However, if it is sufficiently inexpensive, nuclear could be built to serve off-peak loads only, even without the capability to ramp. Doing so would result in the plant having a 100% capacity factor, as it would operate constantly. Moreover, the availability of low-cost CAES could allow even more nuclear capacity to be built to serve on-peak loads. Sioshansi et al. (2012) note that this use of energy storage (*i.e.*, to store nuclear generation overnight to serve on-peak loads) was the major rationale for developing PHS in the United States in the 1970s. The fact that nuclear is not built in any of the cases that we model, even for these limited uses, show that there are substantively lower-cost energy sources, even when seeking carbon-free electricity.

Figure 2.3 summarizes annual expected operating and investment costs of the system with the dif-

ferent carbon-reduction targets. As expected, enforcing and tightening constraint (2.17) increases overall costs. Although the BAU case is the least costly, it is interesting to note that the system can achieve non-trivial carbon-emissions reductions with relatively small overall cost increases relative to BAU. For instance, 25% carbon-emissions reductions can be achieved with only an 8% increase in total expected system costs. On the other hand, complete decarbonization of the electricity sector can be much more costly. 99% emissions reductions increase total expected system costs by 126% (relative to BAU) while complete decarbonization increases costs by 216%.

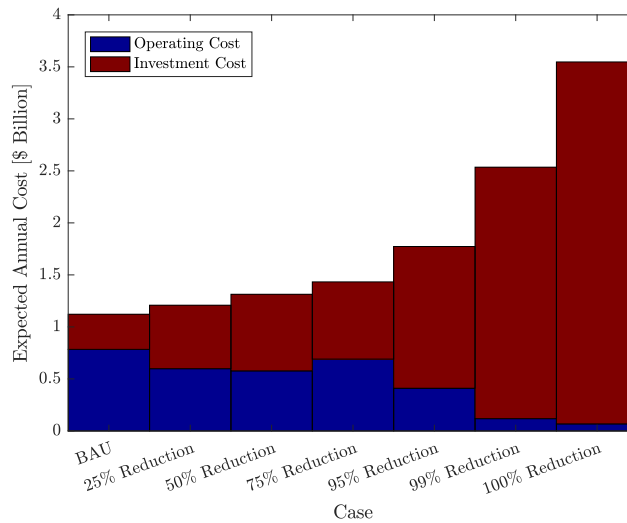


Figure 2.3: Expected Annual Operating and Investment Costs In Business-As-Usual Case and With Different Carbon-Emissions Reductions Relative to Business-As-Usual Case

Table 2.2: Social cost of carbon emission

Carbon emission reduction (%)	Social cost of carbon (\$/t)
75%	\$94/t
85%	\$212/t
95%	\$1,063/t
99%	\$13,167/t
Fully Renewable	\$53,440/t

A question that these cost results raise is what is the optimal level of decarbonization, taking into account its costs and benefits. The United States Environmental Protection Agency (EPA) produced estimates of the social cost of carbon in 2050, under different discount rates and carbon-impact sce-

narios.⁵ We can assess the socially optimal level of carbon reduction by relaxing constraint (2.17) and setting χ (*i.e.*, the carbon-emissions price) equal to a given social-cost-of-carbon estimate in objective function (2.1). Doing so drives the investment model to tradeoff the estimated social cost of carbon against the direct cost of achieving carbon reductions. The EPA estimates a social cost of carbon that ranges between \$69/t and \$212/t. These estimates correspond to socially optimal carbon-reduction levels that range between 25% and 85% (relative to the BAU case). The more restrictive is the CO₂ reduction limit imposed, the higher is the social costs implied. This is confirmed by our analysis showing that a full decarbonization of the considered power system implies a carbon cost of \$53,440/t (see Table 2.2).

2.5 Sensitivity Analyses

This section presents results of six sensitivity analyses, in which we examine the impacts of different case-study assumptions on the design, cost, and operation of the power system. The first sensitivity case considers improvements in the generation, energy-storage, and HVDC technologies (see section 2.5.1). The second examines the impacts of increased transmission congestion (see section 2.5.2). The third examines the impacts of relaxing land-use restrictions on building CSP plants (see section 2.5.3). The fourth sensitivity case analyzes the integration of PV and BES units in the system (see section 2.5.4). The fifth considers adiabatic CAES as an alternative to diabatic CAES (see section 2.5.5). Finally, the sixth evaluates the impacts of demand growth on investment choices (see section 2.5.6). The results of all the sensitivity analyses presented in the following are obtained by assuming no investments in PV and BES units, with the exception of those described in 2.5.4.

2.5.1 Technology Improvement

Our first sensitivity case considers three different technology improvements. The first are efficiency and investment-cost improvements in CSP plants. Lilliestam et al. (2017) examine the historical cost-reduction trajectory of CSP. They find that although CSP has not enjoyed, to date, the same cost-reductions that PV solar has, it is beginning to achieve more aggressive cost reductions.

⁵These estimates, which were produced by the EPA under the administration of President Barack Obama, are archived and publicly available at https://19january2017snapshot.epa.gov/climatechange/social-cost-carbon_.html.

Contemporaneously, Turchi et al. (2013); Neises and Turchi (2014) examine the use of supercritical carbon dioxide power cycles in the power blocks of CSP plants. They find that doing so can achieve powerblock efficiencies that are greater than 50%, as opposed to efficiencies ranging between 35% and 45% for the subcritical Rankine cycles that are used in CSP plants today. Based on these findings, we consider a sensitivity case in which the investment costs of CSP plants are 25% lower than in the base case and their efficiencies (which are used to convert the thermal energy that is gathered by the solar fields into electricity-out capacity factors) are increased to 55%.

The second technology improvement involves offshore wind. International Energy Agency (2016b) finds that offshore wind is costly relative to onshore wind, but that there is potential for this cost difference to be closed. Moreover, offshore wind also suffers from the need for radial HVDC lines to connect generating units to load pockets. Thus, we consider a sensitivity case in which the investment costs of offshore wind plants are reduced 35% (making their costs comparable to onshore wind) and the cost of HVDC lines are reduced an average of 63%, relative to the base case.

The final technology improvement that we consider is a 10% investment-cost reduction for CAES. This reflects the impacts of learning, which can reduce investment costs as CAES plants are developed around the world in the coming years.

Figure 2.4 summarizes the capacity of generating and energy-storage units that are built under the technology-improvement sensitivity cases. The bars in the figure are grouped into four sets of three, each of which corresponds to a different technology-improvement case. The first three sets of bars correspond to one of the three technology improvements (*e.g.*, CSP, offshore wind, or CAES) occurring individually. The fourth set of three bars correspond to a case in which all three of the technology improvements occur simultaneously. For each of the four technology-improvement cases, we examine investment decisions under a BAU case and in cases in which carbon emissions are limited to being 50% and 95% less than in the BAU case.⁶

The figure shows that in all of the technology-improvement cases that we examine, none of the renewable technologies are cost-competitive with thermal generation in a BAU case. This indicates that further technology improvements and cost reductions would be needed to make the renewable technologies that we examine cost-competitive with thermal generation in a BAU case. However, if CAES achieves the investment-cost reductions that we examine here, there is a 15% increase in installed CAES capacity in the BAU case. This increased CAES investment comes with the exact same reduction in CT capacity, meaning that CAES is being used to displace this expensive peaking

⁶The carbon-emissions limits are defined, in all of the sensitivity cases in this section, relative to the carbon emissions in the corresponding BAU case.

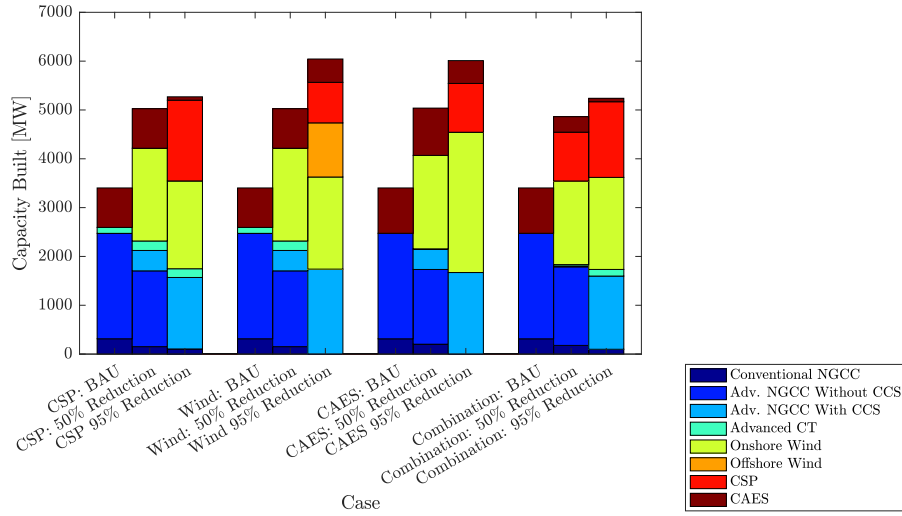


Figure 2.4: Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Technology-Improvement Sensitivity Analysis

capacity.

The technology-improvement cases that we examine result in some changes in the generation mix as carbon-emissions limits are imposed. Reducing carbon emissions by 50% in any of three individual technology-improvement sensitivity cases gives investment mixes that are similar to that in the base case. In all of these cases, there is investment in onshore wind, which displaces thermal generation, to achieve the requisite carbon reductions. However, the investments are different with 50% carbon-emissions reductions in the ‘combined’ technology-improvement sensitivity case. The combined case sees less onshore-wind and CAES investments and greater CSP capacity compared to the three individual technology-improvement sensitivity cases. This is due to the combined impact of higher CSP efficiency and lower CSP- and HVDC-investment costs, which make CSP an economic alternative to onshore wind. Moreover, because CSP has integrated TES, the flexibility that this affords reduces the need for CAES. In sum, the CSP displaces close to 200 MW of onshore wind and 500 MW of thermal capacity, relative to the base case.

These trends continue with a 95%-emissions-reduction target. In all of the sensitivity cases, except for the case in which offshore wind *only* achieves technology improvements, CSP has a greater role in decarbonizing the generation mix. Compared to the base case, the combined and CSP-only technology-improvement sensitivity cases see nearly 1 GW less onshore-wind, 300 MW less CAES, and 600 MW more CSP investments. The fact that 1 GW of wind can be displaced by 600 MW of

CSP capacity further illustrates the flexibility benefit that the TES, which is integrated into the CSP plants, provides. When the system relies on wind for decarbonization in the base case, greater capacity must be built to manage the variability in real-time wind availability. The TES systems in the CSP plants reduce the need for such oversizing, as the TES can be used to mitigate the impacts of variable solar availability.

The sensitivity case in which offshore wind *only* has technology improvements results in 1.1 GW of offshore wind being built in place of about 200 MW of CSP (relative to the base case) to achieve 95% emissions reductions. This can be contrasted with the base case, in which offshore wind is only built to achieve complete decarbonization, due to its relatively high cost. Interestingly, no offshore wind is built to achieve 95% emissions reductions in the combined technology-improvement case. This is because, although offshore wind is economically viable in the combined technology-improvement case, so too is CSP, which is less costly and more flexible than offshore wind.

Figure 2.5 summarizes the annual expected operating and investment costs of the system under the four technology-improvement sensitivity cases. The figure shows that these sensitivity cases yield mixed cost savings. For instance, the combined technology-improvement cases reduces the cost of achieving 50% and 95% emissions reductions by 22% and 33%, respectively, relative to the base case. The case in which CSP *only* has technology improvements reduces the cost of achieving these emissions reductions by 6% and 22%. This can be contrasted with a case in which offshore wind *only* has technology improvements, which only yields an 8% reduction in the cost of achieving 95% emissions reductions. This follows from the results that are summarized in Figure 2.4—the combined and CSP-only technology-improvement sensitivity cases yielded the most notable differences in the generation mix.

2.5.2 Transmission Congestion

Our second sensitivity analysis considers the impacts of diminished transmission capacity, which can result from smaller-sized transmission equipment being installed and land-use or other restrictions on building candidate lines. We consider a case in which the transmission capacity of all existing and candidate lines is reduced by 25%. We further assume that five of the candidate AC lines, which are close to load pockets in the transmission network, are no longer available to be built.

Figure 2.6 shows the mix of generation and CAES units that are built in the transmission-congestion sensitivity case with different emissions-reduction targets. Overall, the transmission-congestion

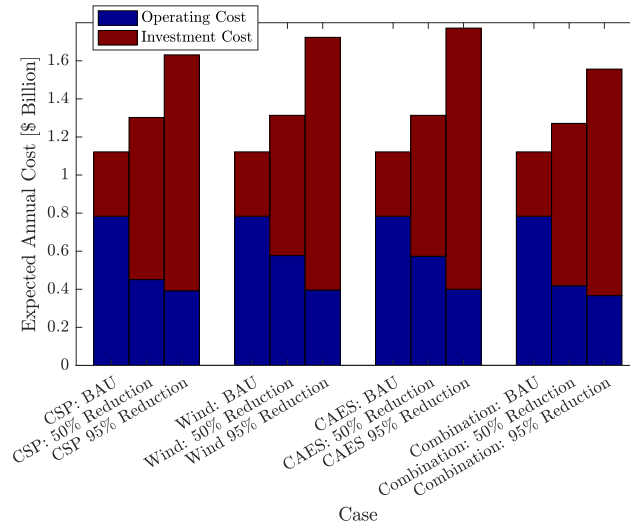


Figure 2.5: Expected Annual Operating and Investment Costs In Business-As-Usual Case and With Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Technology-Improvement Sensitivity Analysis

sensitivity case results in greater transmission investments and requires more generation and CAES capacity to be built, relative to the base case. The base case does not require new transmission builds until reducing carbon emissions by 95%. Moreover, these transmission builds in the base case are primarily to deliver energy from remote renewable units to loads. Conversely, the transmission-congested sensitivity case requires transmissions builds even in the BAU case. These transmission builds are needed simply to avoid load curtailment. The 99%-carbon-reduction case requires more than double the transmission investment relative to the base case.

The generation and CAES units are more dispersed throughout the transmission network than in the base case. Dispersing the energy-production and -storage capacities across the network provides greater flexibility in having energy available in congested load pockets. One benefit of renewable units in this regard is that they are fairly dispersed within the network. As such, renewable and CAES units are jointly able to provide a great deal of congestion relief. One interesting observation regarding the operation of the system in the transmission-constrained case is that some hours see simultaneous charging and discharging of energy-storage units at different locations in the network. This is inefficient, insomuch as energy is lost when it is stored. Simultaneous charging and discharging of energy storage is nonetheless necessary, because load peaks and troughs occur at different times at different locations. As such, absent transmission capacity, excess energy must be stored at some nodes in the network despite stored energy being discharged to serve load

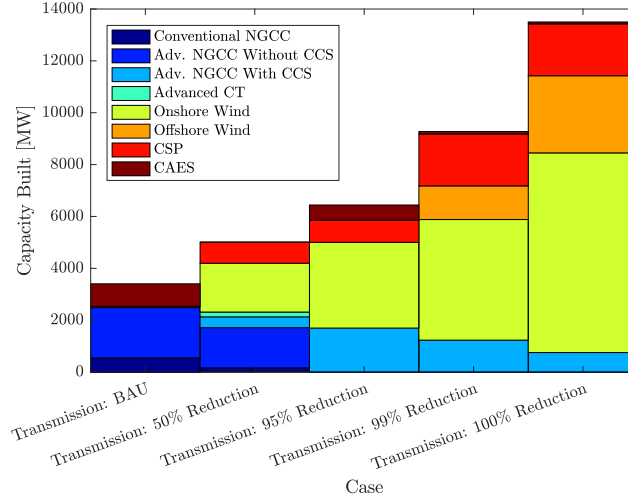


Figure 2.6: Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Transmission-Congestion Sensitivity Analysis

elsewhere in the network.

We also evaluate the role played by the transmission system in abating CO₂ emissions. To this aim, we assume a social cost of carbon equal to \$100/t, which is between the values of \$69/t and \$212/t estimated by EPA (see section 2.4), and we compare the total amount of emissions generated in the base case with that obtained when the transmission capacity of the system is downgraded. Interestingly, there is a slight reduction of carbon emissions (−0.22%) when lines capacity decreases, although the total cost is higher (+2.3%). This is due to the specific network configuration, which accounts for several candidate nodes for the construction of renewable and storage facilities but only a few for thermal plants. As a consequence, the transmission downgrade implies that thermal capacity is partially offset by both additional wind capacity (+141 MW) and storage (CAES) capacity (+110 MW) which allows to overcome the congestion. Since CAES and wind plants are less pollutant than thermal plants, the overall carbon emission level decreases but the total costs increase driven by the higher investment costs.

2.5.3 Land Use

The purpose of the land-use sensitivity case is to further explore the economic competitiveness of CSP relative to onshore wind. Our base case places CSP at a competitive disadvantage compared to onshore wind, because CSP can only be connected to the network via radial HVDC transmission

lines. As such, CSP bears an additional cost compared to onshore wind, which can use existing transmission corridors and can, in some cases, be co-located with load. The land-use sensitivity case explores the competitiveness of CSP independent of transmission-interconnection costs, by allowing CSP plants to be built at each node where onshore wind can be. These additional candidate CSP plants are assumed to have the same cost and technical characteristics as the CSP units in the base case. We use weather data for 18 additional locations, which correspond to the nodes where onshore wind can be built, to simulate real-time availability of the added candidate CSP plants. We also consider a sensitivity case, which removes the land-use restrictions on building CSP plants and which assumes that CSP achieves the technology improvements that are outlined in section 2.5.1. Figure 2.7 summarizes the generation and energy-storage capacities that are built with different emissions-reductions targets under the two land-use sensitivity cases. The first set of six bars corresponds to a case in which *only* land-use restrictions on the deployment of CSP are relaxed. The second set of six bars corresponds to a case in which land-use restrictions are relaxed and the CSP-technology improvements that are outlined in section 2.5.1 are assumed. The first set of bars shows that even without the cost of interconnecting with the transmission network, CSP is not cost competitive with onshore-wind units. Indeed, the only difference in CSP investments that results from relaxing land-use restrictions is that slightly more capacity is built to achieve 99% emissions reductions. In fact, we find that with land-use restrictions CSP is only built when the capacity to build onshore wind at a particular node is exhausted.

The second set of six bars in Figure 2.7 shows that this result is reversed if land-use restrictions are relaxed *and* CSP achieves technology improvements. Considerably more CSP capacity is built under these assumptions compared to the base case (and the land-use-*only* sensitivity case) to achieve carbon-emissions reductions between 25% and 95%. Interestingly, the amount of CSP capacity that is built to achieve 99% carbon-emissions reductions is the same with and without the CSP-technology improvements. This suggests that with 99%-carbon-emissions reductions, the CSP that is built is solely driven by the emissions constraint and not by the economics of the technology. The findings of this sensitivity case are consistent with current experience in the deployment of renewable energy sources. Although CSP has a flexibility benefit relative to wind (and PV solar) that stems from the ability to integrate low-cost TES, CSP has seen little deployment compared to wind. This is largely because CSP is considerably more expensive to build (even when neglecting transmission-interconnection costs). However, our sensitivity cases in which CSP achieves technology improvements (both the case examined here and those in section 2.5.1) reveal that CSP has the potential to become highly competitive with wind.

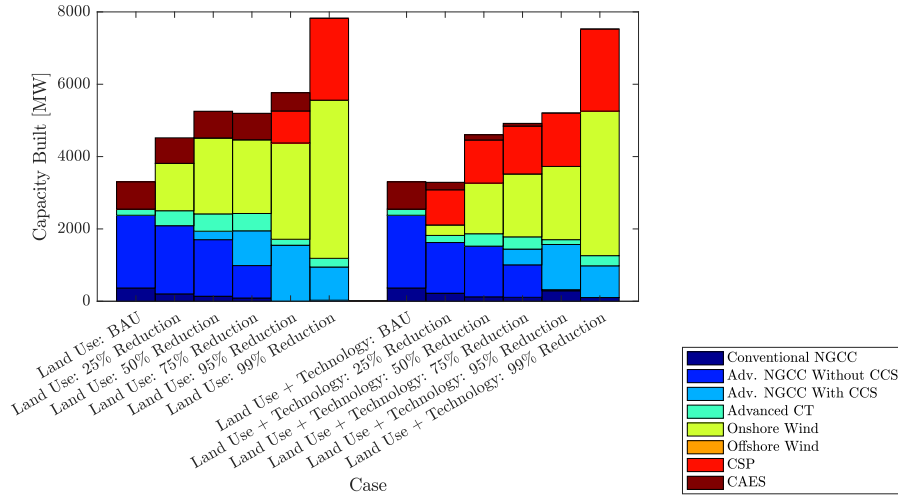


Figure 2.7: Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Land-Use Sensitivity Analysis

2.5.4 PV solar and BES

This sensitivity analysis examines the benefits of adding PV solar and BES, the characteristics of which are detailed in the , to the set of technologies that are available. As described in International Renewable Energy Agency (2018b) the global PV market has grown exponentially in the last decade and the technology experienced a huge cost reduction. Moreover, the declining costs of energy storage technologies have stimulated interest in combining PV with BES units to provide dispatchable energy and reliable capacity (Feldman and Margolis (2018)). In particular, Fu et al. (2018) show that the use of lithium-ion batteries in U.S. utility-scale applications has grown in recent years owing to the technology's favorable cost and performance characteristics.

Figure 2.8 summarizes the capacity of generating and energy-storage units that are built when PV solar and BES are added to the set of technologies that can be installed.

Interestingly, the figure shows that PV units are cost-competitive with thermal generation in the BAU case. Without any carbon policy in-place in the BAU case, 2.3 GW of PV solar is built to offset 230 MW of thermal capacity and 35 MW of CAES relative to the base case that is described in Section 2.4. This configuration with PV solar is beneficial from environmental and economic perspectives. CO₂ emissions and total costs are 25% and 20% lower, respectively, due to PV solar reducing operating costs by 25% with only a 3% increase in investment costs. 770 MW of CAES is still built in the BAU case to serve on-peak demand, because using CAES for peaking purposes is less costly

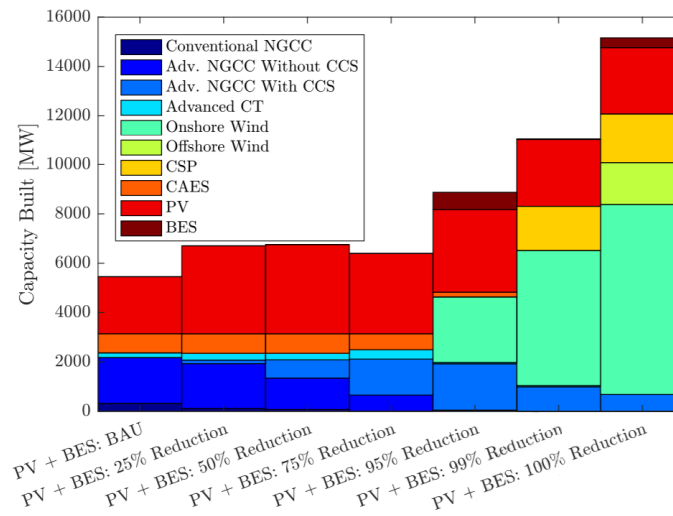


Figure 2.8: Capacity built in business-as-usual case and with different carbon-emissions reductions relative to business-as-usual case under PV-solar-and-BES sensitivity analysis

than using BES is. As the carbon-emissions limit is enforced and tightened, the thermal-generation mix shifts toward the use of CCS and PV-solar capacity increases. Carbon-emissions reductions beyond 75% require the use of onshore wind at nodes at which PV cannot be built (or at which all available PV capacity is built). 2.5 GW of onshore wind is built to achieve 95% carbon reductions and 700 MW of BES capacity replaces almost all CAES capacity. Achieving 99% carbon emissions reductions shows another interesting result: PV-solar and BES capacities both decrease in favor of CSP. Given their relative costs and efficiencies, the TES in CSP provides the preferred source of dispatchability and flexibility to achieve these extreme carbon reductions. This is consistent with the findings of Madaeni et al. (2012). To achieve complete decarbonization, all available CSP capacity is built, more BES capacity is built, and offshore wind enters the generation mix.

2.5.5 Adiabatic CAES

This sensitivity case considers the use of adiabatic CAES as an alternative to diabatic CAES. Unlike diabatic CAES, adiabatic CAES is a pure energy-storage technology. The primary difference between diabatic- and adiabatic-CAES systems is the treatment of the waste heat that is produced when air is compressed in the storage cycle. In a diabatic-CAES system, the waste heat is exhausted to the atmosphere. As such, a diabatic-CAES system must combust natural gas as a heat source when expanding the compressed air in the discharging cycle. An adiabatic-CAES system, con-

versely, stores the waste heat (often in a dedicated TES system). This stored heat is then combined with the compressed air during the discharging cycle, alleviating the need to use natural gas. This makes adiabatic CAES considerably less costly to operate than diabatic CAES. Moreover, adiabatic CAES involves no direct carbon emissions, unlike a diabatic-CAES system, which emits carbon when combusting the natural gas in the discharging cycle. Hartmann et al. (2012); Barnes et al. (2015); Barbour et al. (2015); Liu and Wang (2016); Sciacovelli et al. (2017) provide further details regarding the technical properties and viability of adiabatic CAES.

The adiabatic-CAES sensitivity case assumes that this technology is available in place of diabatic CAES. It is difficult to estimate the technical characteristics of adiabatic CAES, because no commercial-scale systems exist today. Indeed, the DOE Global Energy Storage Database⁷ lists the 500 kW Pollegio-Loderio Tunnel Demonstration Plant in Switzerland as the only operational adiabatic CAES plant as of July, 2018. Thus, we rely on current and future estimates of the capabilities of these plants in modeling the adiabatic-CAES sensitivity case.

We assume that the adiabatic-CAES plants have the $\bar{c}_s = 10$ hours of storage capacity and that each unit can have up to 350 MW of power capacity built. The adiabatic-CAES units are assumed to have efficiencies of $\eta_s^S = 1.33$, which means that 1.33 MWh of energy must be stored to later discharge 1 MWh of electricity. η_s^S being greater than 1 reflects the fact that adiabatic CAES is a pure energy-storage technology, meaning that electricity is lost in the storage cycle. However, the benefit of this energy loss is that the adiabatic-CAES units have zero carbon emissions.

We consider two sets of investment and operating costs for the adiabatic-CAES plants. The first set, which reflect current cost estimates, assumes investment costs ranging between \$1187/kW and \$1210/kW and charging and discharging costs of \$14/MWh and \$17/MWh, respectively. The second set, which reflect possible future cost reductions, assumes investment costs ranging between \$594/kW and \$605/kW and charging and discharging costs of \$12/MWh and \$13/MWh, respectively.

Figure 2.9 summarizes the generation and energy-storage capacities that are built in the adiabatic-CAES sensitivity case with different emissions-reduction targets. The first set of bars corresponds to a case with high investment and operating costs for adiabatic CAES, whereas the second set corresponds to lower future cost projections. Interestingly, we find that adiabatic CAES is a considerably less desirable energy-storage technology compared to diabatic CAES. The base case and other sensitivity cases result in some diabatic CAES being built, except with extremely high emissions-reduction targets. This is because once the emissions-reduction target is sufficiently stringent, the

⁷<https://www.energystorageexchange.org/>

carbon emissions that are associated with discharging diabatic CAES makes it use infeasible.

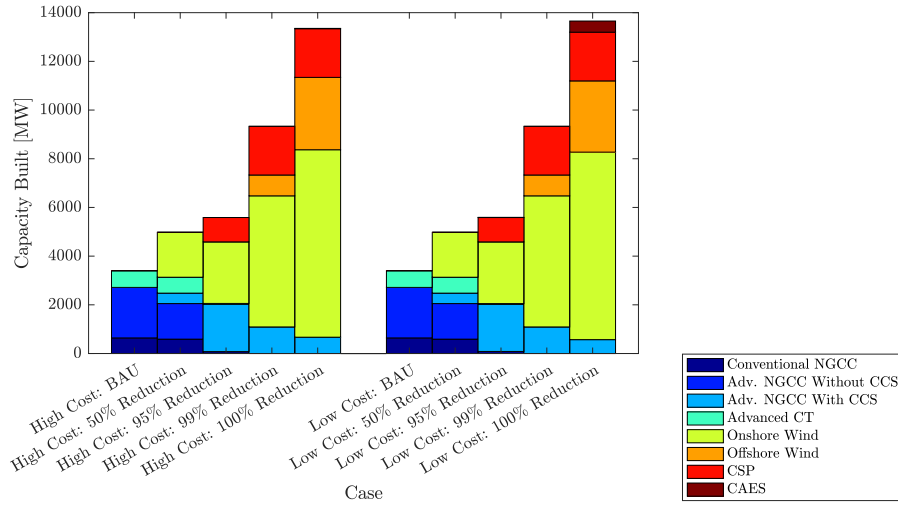


Figure 2.9: Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Adiabatic-CAES Sensitivity Analysis

Adiabatic CAES relaxes the impact of the emissions-reduction constraint on the use of CAES. However, adiabatic CAES is more costly than diabatic CAES. This cost disadvantage overwhelms the zero-emissions benefit of adiabatic CAES, except in the extreme case of full decarbonization. With full decarbonization, about 10 MW and 500 MW of adiabatic CAES is built with high and low costs, respectively, as opposed to no diabatic CAES in the base case.

2.5.6 Electricity demand growth

According to Mai et al. (2018), electrification and the adoption of electric vehicles may cause electricity demand in the United States to increase over the next three decades. Cooper and Schefter (2017) foresee an additional seven million electric vehicles by 2025 in the United States, compared to the 567,000 electric vehicles as of the end of 2016. Rissman (2017) estimates that electric vehicles will compose 60%–75% of total new light-duty vehicle sales in the United States by 2050, at which point they will represent 13%–15% of total national electricity demand. Given this potential for substantive electricity-demand increases, we conduct a sensitivity analysis on demand growth by modifying the demand levels in the second-stage scenarios. Specifically, we examine three cases in which electricity demands increase by 10%, 30%, and 50% in the high-demand scenario (as opposed to the 5% demand growth that is assumed in the base case that is presented in Sections 2.3 and 2.4).

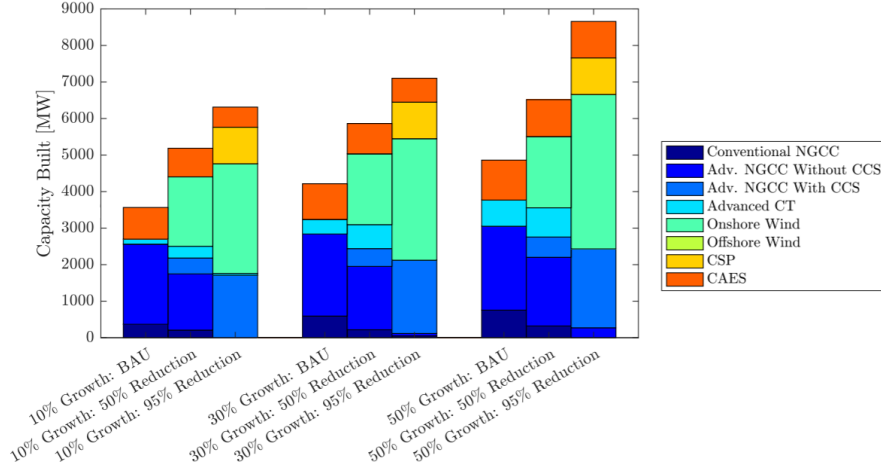


Figure 2.10: Capacity built in business-as-usual case and with different carbon-emissions reductions relative to business-as-usual case under electricity-demand sensitivity analysis.

The no-demand growth and low-demand scenarios are not modified in this sensitivity analysis relative to the base case. Fig. 2.10 shows generation and energy-storage units that are built in the electricity-demand sensitivity case with different emissionsreductions targets. Contrasting this with Fig. 2.2 shows that regardless of the demand-growth level in the high-demand scenario, the technology mix that is built is not altered fundamentally. Instead, the sole substantive impact of demand growth is to “rescale” the total capacity that is built to serve the possible peak load in the high-demand scenario. This increased capacity translated into higher system costs, which are shown in Fig. 2.11.

2.6 Discussion and Conclusions

This Chapter presents a two-stage stochastic optimization model that can be used for long-term power system-expansion planning.

We pay particular attention to modeling carbon-emissions limits and decarbonization of electricity production. This can be achieved in our model using two policy mechanisms: either via an explicit carbon-emissions constraint or by pricing carbon (e.g., Pigouvian taxes on emissions). Our model and case study consider three technical options for decarbonization – renewable energy sources, nuclear generators, and fossil-fueled plants with integrated CCS systems. Weather-dependent renewable generators raise modeling, planning, and operating challenges, because their real-time

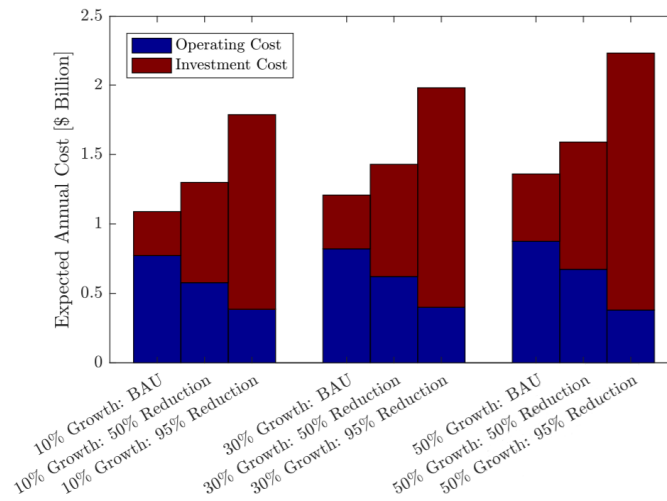


Figure 2.11: Expected annual operating and investment costs in business-as-usual case and with different carbon-emissions reductions relative to business-as-usual case under electricity-demand sensitivity analysis.

resource availability is uncertain and variable. As such, we include energystorage technologies in the model as a flexibility source for mitigating these characteristics of renewables. We focus on two energystorage technologies—CAES and TES integrated in CSP plants—in our base case, which can be deployed widely in high-renewable penetration scenarios. The real-time variability of renewable availability is captured in our model using a variety of different operating conditions. These operating conditions capture the range of weather conditions over the course of a year, which can give different load levels and wind and solar availabilities. Large-scale uncertainties, such as long-term load growth, are captured explicitly through second-stage scenarios in the scenario tree. We demonstrate our model and analyze decarbonization pathways using a detailed case study that is based on the ERCOT system. We find that absent any policy mechanism that internalizes the societal cost of carbon, the power system still maintains a fossil-fueled generation mix with some CAES to alleviate the need to build peaking generation capacity. Natural gas is preferred to coal as a generation fuel, given its anticipated low cost and the high efficiency of NGCC units. This system design holds regardless of potential future cost reductions in renewable and CAES technologies. However, when investments in PV solar and BES units are allowed, 2.3 GW of PV-solar capacity is installed and displaces 230 MW of thermal and 35 MW of CAES in the BAU case. This provides both economic and environmental benefits and shows that PV solar is indeed cost-competitive with thermal generation absent any carbon policy.

As the social cost of carbon is internalized (we do this in our case study via explicit carbon-emissions limits) the system moves away from natural gas-fired generators toward onshore-wind and PV solar (when allowed) capacity. Moreover, portions of the natural gas-fired fleet uses CCS for modest carbon-emissions limits. As the carbon-emissions limits become very stringent, fossil-fueled generation essentially is phased out of the system completely. CAES is also phased out, because of the carbon emissions that are associated with the combustion of natural gas in the discharging cycle. The flexibility that CAES affords is instead provided by CSP plants, which have TES incorporated in them. Finally, achieving complete decarbonization of the electric power system calls upon the use of offshore wind, which tends to be more costly than onshore with only marginal relative improvements in its capacity factor. CSP plants are preferred to PV solar and BES to achieve extreme decarbonization. Interestingly, nuclear is not used for decarbonization purposes, given its very high cost.

Achieving modest decarbonization targets yields relatively small expected power system-cost increases. Indeed, we find that for a wide range of social costs of carbon, reducing carbon emissions by between 25% and 85% (relative to a BAU case) is socially optimal. One of the challenges in setting carbon policy is that the social costs of climate change are difficult to estimate a priori. Thus, it may prove robust against the potential for catastrophic impacts of climate change to set decarbonization targets that are toward the upper end of this range.

Under current estimates of costs and technical characteristics, onshore wind and solar PV are the most attractive renewable-energy technologies that we consider. We find, however, that if CSP achieves its projected cost reductions and efficiency gains and if transmission interconnection costs are reduced, CSP can become an extremely competitive alternative to onshore wind. Moreover, the inherent flexibility of CSP alleviates the need to build stand-alone energy storage capacity. Offshore wind is not competitive with onshore wind and CSP. This is because of its relatively high investment and operating costs and the need to build radial HVDC lines to connect it to load centers. For these reasons, offshore wind is built only if all other renewable-energy sources are exhausted.

Diabatic CAES is an economically feasible alternative to PHS. This is an important finding, because many parts of the world lack either the water or geological formations that are necessary for the deployment of PHS. Conversely, CAES has the potential to be deployed in many renewable-rich regions. The major difficulty that diabatic CAES raises is the emission of carbon that is associated with the discharging cycle. Adiabatic CAES, which is a technology that is currently in the demonstration phase, holds the potential to mitigate the carbon-emissions issue of diabatic CAES. However, our adiabatic-CAES sensitivity case reveals that with current and future projections of

technology costs, adiabatic CAES is not an economic source of supply-side flexibility.

Our findings suggest that adiabatic CAES requires further development to become a viable alternative to diabatic CAES, except under extreme carbon-reduction targets. Our analysis focused on using energy storage as a source of supply-side flexibility to achieve supply/demand balance. One could also consider the use of demand-side flexibility, for instance through demand response, as a means of achieving this. The use of demand response is an area of future research that is beyond the scope of our current work. The methodology that Sioshansi and Denholm (2010a,b) uses to examine the renewable-integration benefits of demand response likely could be adapted within the modeling framework that we develop here.

Our analysis shows that demand growth does not affect the optimal mix of technologies that are built. Rather, demand growth results in scaling up the total capacity that is installed with a commensurate increase in system costs.

Appendix

This appendix provides all of the detailed input data that are used in the base-case and sensitivity analyses in the Chapter “A Two-Stage Stochastic Optimization Planning Framework to Deeply Decarbonize Electric Power Systems.” All of the investment costs are overnight costs, which must be annualized (we use an 11% annualization rate in our analysis).

2.A Data For Base Case

Table 2.A.1: Technical Characteristics of Onshore-Wind Units

Node	Maximum Capacity [MW]	Operating Cost [\$/MWh]	Investment Cost [\$/kW]
1	700	2.71	2005
7	700	2.15	2022
13	700	2.06	1980
13	700	2.05	2010
14	700	2.10	1990
14	700	2.21	2020
16	700	2.04	2040
18	700	2.00	2100
21	700	2.04	2000
22	700	2.55	1970
23	700	2.02	1980

Table 2.A.2: Technical Characteristics of Offshore-Wind Units

Node	Maximum Capacity [MW]	Operating Cost [\$ /MWh]	Investment Cost [\$ /kW]
25	2000	4.85	3190
27	2000	5.35	2900

Table 2.A.3: Technical Characteristics of CSP Units

Node	Maximum Capacity [MW]	Generation Cost [\$ /MWh]	TES-Discharging Cost [\$ /MWh]	TES-Charging Cost [\$ /MWh]	Investment Cost [\$ /kW]
26	1000	4.00	0.1400	0.1549	3272
28	1000	3.80	0.1400	0.1549	3599

Table 2.A.4: Technical Characteristics of Diabatic-CAES Units

Node	Maximum Capacity [MW]	Discharging Cost [\$ /MWh]	Charging Cost [\$ /MWh]	Investment Cost [\$ /kW]	Carbon-Emissions Rate of Discharging [t /MWh]
1	350	28.79	0.5	571	0.259
7	350	28.71	0.5	560	0.259
13	350	28.33	0.5	599	0.259
14	350	28.47	0.5	550	0.259
16	350	28.55	0.5	605	0.259
21	350	28.66	0.5	610	0.259
22	350	28.75	0.5	589	0.259
23	350	28.41	0.5	600	0.259

Table 2.A.5: Technical Characteristics of Existing Transmission Lines

From Node	To Node	Reactance [p.u.]	Flow Limit [MW]
1	2	0.014	450
1	3	0.211	220
1	5	0.085	510
1	23	0.087	600
2	4	0.127	220
2	6	0.192	220
3	9	0.119	400
3	24	0.084	400
4	9	0.104	220
5	10	0.088	220
6	10	0.061	200
7	8	0.061	320
8	9	0.165	220
8	10	0.165	220
9	11	0.084	320
9	12	0.084	250
10	11	0.084	230
10	12	0.084	200
11	13	0.048	400
11	14	0.042	320
12	13	0.048	400
12	23	0.097	600
14	16	0.059	600
15	16	0.071	600
15	21	0.049	600
15	21	0.049	600
15	24	0.052	310
16	17	0.026	800
16	19	0.023	600
17	18	0.014	600
17	22	0.105	600
18	21	0.026	600
18	21	0.026	600
19	20	0.040	600
19	20	0.040	600
20	23	0.022	600
20	23	0.022	600
21	22	0.068	600

Table 2.A.6: Technical Characteristics of Candidate ac Transmission Lines

From Node	To Node	Reactance [p.u.]	Flow Limit [MW]	Investment Cost [\$ million]
1	9	0.026	600	210
3	15	0.026	600	200
3	18	0.049	400	312
7	10	0.026	600	201
9	14	0.026	600	190
9	21	0.049	400	315
10	11	0.026	600	180
10	13	0.026	600	195
12	22	0.049	400	290
16	17	0.026	800	190

Table 2.A.7: Technical Characteristics of Candidate HVDC Transmission Lines

From Node	To Node	Flow Limit [MW]	Investment Cost [\$ million]
25	6	2000	684.0
25	8	2000	817.5
25	13	2000	589.5
25	22	2000	503.5
26	3	1000	503.5
26	15	1000	551.0
26	16	1000	589.0
26	17	1000	617.5
27	6	2000	532.0
27	8	2000	551.0
27	13	2000	503.5
28	1	1000	817.5
28	3	1000	684.0
28	15	1000	589.5
28	18	2000	503.5

2.B Data For Sensitivity Cases

2.B.1 Data For Technology-Improvement Sensitivity Case

Table 2.B.1: Investment Costs of Candidate HVDC Transmission Lines

From Node	To Node	Investment Cost [\$ million]
25	6	304.0
25	8	437.0
25	13	209.0
25	22	123.5
26	3	123.5
26	15	171.0
26	16	209.0
26	17	237.5
27	6	152.0
27	8	171.0
27	13	123.5
28	1	437.5
28	3	304.5
28	15	209.0
28	18	123.5

2.B.2 Data For Land-Use Sensitivity Case

Table 2.B.2: Technical Characteristics of Additional CSP Units in Land-Use Sensitivity Case

Node	Maximum Capacity [MW]	Generation Cost [\$/MWh]	TES-Discharging Cost [\$/MWh]	TES-Charging Cost [\$/MWh]	Investment Cost [\$/kW]
1	350	3.86	0.14	0.1549	3507
7	350	3.87	0.14	0.1549	3442
13	350	3.86	0.14	0.1549	3336
13	350	3.81	0.14	0.1549	3471
14	350	3.88	0.14	0.1549	3513
14	350	3.85	0.14	0.1549	3350
16	350	3.87	0.14	0.1549	3239
18	350	3.85	0.14	0.1549	3345
21	350	3.82	0.14	0.1549	3432
22	350	3.83	0.14	0.1549	3222
23	350	3.89	0.14	0.1549	3417
26	1000	4.00	0.14	0.1549	3599
28	1000	3.80	0.14	0.1549	3272

Table 2.B.3: Technical Characteristics of Additional CSP Units in Land-Use Sensitivity Case With Technology Improvements

Node	Maximum Capacity [MW]	Generation Cost [\$/MWh]	TES-Discharging Cost [\$/MWh]	TES-Charging Cost [\$/MWh]	Investment Cost [\$/kW]
1	350	3.86	0.14	0.1549	2630
7	350	3.87	0.14	0.1549	2581
13	350	3.86	0.14	0.1549	2502
13	350	3.81	0.14	0.1549	2603
14	350	3.88	0.14	0.1549	2634
14	350	3.85	0.14	0.1549	2512
16	350	3.87	0.14	0.1549	2429
18	350	3.85	0.14	0.1549	2508
21	350	3.82	0.14	0.1549	2574
22	350	3.83	0.14	0.1549	2416
23	350	3.89	0.14	0.1549	2562
26	1000	4.00	0.14	0.1549	2699
28	1000	3.8	0.14	0.1549	2454

2.B.3 Data For Adiabatic-CAES Sensitivity Case

Table 2.B.4: Technical Characteristics of Adiabatic-CAES Units With High Costs

Node	Maximum Capacity [MW]	Discharging Cost [\$/MWh]	Charging Cost [\$/MWh]	Investment Cost [\$/kW]
1	350	14	17	1198
7	350	14	17	1187
13	350	14	17	1199
14	350	14	17	1210
16	350	14	17	1195
21	350	14	17	1205
22	350	14	17	1189
23	350	14	17	1200

Table 2.B.5: Technical Characteristics of Adiabatic-CAES Units With Low Costs

Node	Maximum Capacity [MW]	Discharging Cost [\$/MWh]	Charging Cost [\$/MWh]	Investment Cost [\$/kW]
1	350	11.5	12.5	599
7	350	11.5	12.5	593
13	350	11.5	12.5	599
14	350	11.5	12.5	605
16	350	11.5	12.5	597
21	350	11.5	12.5	602
22	350	11.5	12.5	504
23	350	11.5	12.5	600

2.B.4 Data For PV and BES Sensitivity Case

Table 2.B.6: Technical Characteristics of BES Units

Node	Maximum Capacity [MW]	Discharging Cost [\$/MWh]	Charging Cost [\$/MWh]	Investment Cost [\$/kW]
18	350	1	0.5	2000
16	350	1	0.5	2090
3	350	1	0.5	1960
1	350	1	0.5	1990
14	350	1	0.5	2100
5	350	1	0.5	1971
22	350	1	0.5	1955
23	350	1	0.5	2010
10	350	1	0.5	2012
7	350	1	0.5	1989

Table 2.B.7: Technical Characteristics of PV Units

Node	Maximum Capacity [MW]	Operating Cost [\$/MWh]	Investment Cost [\$/kW]
18	700	1.5	790
16	700	1.51	801
3	700	1.52	1000
1	700	1.53	799
14	700	1.511	798
5	700	1.521	1010
22	700	1.531	800
23	700	1.54	792
10	700	1.52	1012
7	700	1.521	795

Chapter 3

Robust expansion planning of a distribution system with electric vehicles, storage and renewable units

The decarbonization of energy systems, as foreseen by the European Commission, passes through the transition towards low- and zero-emission vehicles and the investments in efficient technologies. To this end this chapter proposes an adaptive robust optimization approach for the expansion planning problem of a distribution system where expansion decisions involve the construction of candidate renewable generating units, storage units, and charging stations for electric vehicles. The problem is formulated under the perspective of a central planner that aims at determining the expansion plan that minimizes both investment and operation costs, including the power exchanged with the transmission system. Both short-term variability and long-term uncertainty are considered in the proposed approach and are modeled in different ways. In particular, short-term variability of the demand, the production of stochastic units, and the price of electricity withdrawn from or injected into the transmission system is modeled using a number of representative days corresponding to different operating conditions. Long-term uncertainty in the future peak demands, the future value of electricity exchanged with the transmission grid, and the number of electric vehicles is instead modeled through confidence bounds. A case study based on a 69-node distribution network is used to illustrate the effectiveness of the proposed technique, as well as to analyze the relationship between the optimal expansions decisions, the revenues from selling electricity to the electric vehicles, the degree of independence from the transmission system, and the

role played by the investment budget availability. Moreover, an ex-post decarbonization analysis is conducted to evaluate the environmental impact of the adoption of electric vehicles. Finally, the proposed approach is compared with a stochastic model in terms of computational performance. The model developed in this chapter have been implemented in GAMS and solved with CPLEX.

3.1 Introduction

The implementation of the Energy Roadmap 2050 (see European Commission (2011a)) commits Europe to reducing greenhouse gas (GHG) emissions to 80-95% below 1990 levels by 2050. This target can be accomplished thanks to the full decarbonization of energy systems, the transition towards low- and zero-emission vehicles, and the investments in efficient technologies. This is a direct consequence of the fact that the energy sector is the main responsible of carbon emissions, followed by the transport sector, which generates around a quarter of the European GHG emissions. It is clear that, in the future, a large share of the planned GHG emissions reductions should come from road transport. However, while GHG emissions from energy and industrial sectors have been reduced in recent decades, those from transport have significantly increased. For instance, emissions from road transport are today around 17% above 1990 levels.

For this reason, the European Commission outlined in a White Paper issued in 2011 (see European Commission (2011b)) a roadmap for the transport sector to achieve a 60% reduction in its GHG emission levels compared with those of 1990 by 2050.

This White Paper shows how the transition to a more sustainable transport system can be addressed by reducing the European dependence on oil leading both to a reduction of GHG emissions and to an improvement of energy security. In particular, it describes goals for a competitive and efficient transport system, including benchmarks, such as halving the utilization of conventionally fuel cars in urban transport by 2030 and phasing them out entirely by 2050. More recently, the European Commission has published a guideline for low-emission mobility that highlights the importance of removing obstacles to the transport electrification with the aim of moving towards low- and zero-emission vehicles and increasing the supply of low-emission alternative fuels, such as renewable energy (see European Commission (2016)).

The direct consequence of the implementation of these policies put forward by the European Commission is a progressively increasing penetration of electric vehicles (EVs) in the coming years. This may impose additional challenges to distribution systems, which are already stressed by the uncertainty and variability of renewable generating units.

This chapter focuses on these aspects and proposes a static expansion planning model for a distribution system in which investment decisions are made for a future target year. This assumption is customary in the technical literature (see, e.g., Baringo and Baringo (2016); Baringo and Baringo (2017); Kazempour et al. (2013); Moreno et al. (2017); Wang et al. (2009); Zhang and Conejo (2017)). The analysis is done taking the perspective of a central planner that can invest in storage units and photovoltaic (PV) solar power plants in addition to charging stations for EVs. Storage units are used to compensate for the variability of the electricity production from PV power plants. No existing power capacity is assumed and therefore demands can be supplied by the electricity either produced by PV power plants, or provided by storage units, or withdrawn from the transmission grid (main grid in the following) with which the distribution system is connected. Notice that the distribution system is allowed to sell electricity to the main grid in the case where its electricity production exceeds the local needs.

These storage units, PV power plants, and charging stations (SRC) investment decisions are taken on the basis of both short-term variability and long-term uncertainty. Short-term variability describes the daily intermittent production of the stochastic PV units, as well as the daily variation of the residential demand, the driving requirements of EV owners, and the price of electricity exchanged with the main grid; while long-term uncertainty regards the future peak demand, the future price of the electricity exchanged with the main grid, and the future number of EVs. The short-term variability is modeled through representative days in the form of operating conditions (see Section 3.2.2).

In contrast, two methods are generally considered in the technical literature to deal with medium- and long-term uncertainty in this type of decision-making problems. One of them is stochastic programming (see, Birge and Louveaux (1997)) that models uncertain parameters using a set of scenarios. The main drawback of this approach is that its accuracy strongly depends on the knowledge of the probability distribution of uncertain parameters and on the number of scenarios. In particular, for a long-term planning problem such as the one considered in this chapter, it is not straightforward to have a good forecast of probability distributions of uncertain parameters. Moreover, increasing the number of scenarios may result in an intractable problem. The other option used in the technical literature is robust optimization (see Bertsimas and Sim (2004)) that ensures worst case protection within confidence bounds. This approach for modeling decision-making under uncertainty does not require scenarios since uncertain parameters are expressed through uncertainty sets. This makes robust optimization problems more tractable than stochastic programming models from a computational point of view. Notice that a robust solution must be feasible for any

realization of the uncertainty, and optimal in the worst-case choice of the stochastic parameters in the uncertainty set. Adaptive Robust Optimization (ARO) gives the possibility to model decision making under recourse (see Bertsimas and Brown (2009); Bertsimas et al. (2011)) and represents the correspondent of stochastic programming with recourse. The ARO approach aims at minimizing the total cost in the worst-case realization of the uncertain parameters, assuming that recourse actions can be taken as a response to the uncertainty (see Yanıkoğlu et al. (2019), for a review). This leads to the formulation of a three-level optimization problem that can be summarized as follows:

- In the first level, expansion and operating costs are minimized considering as variables those associated with the investments or expansion decisions.
- The second level identifies the worst realization of the (long-term) uncertain parameters in the considered confidence bounds through the maximization of the operating costs of the system. The variables of this level are those modeling the uncertain parameters.
- In the third level, the operating costs are minimized taking as given the values of the first and the second level variables.

Recent ARO applications on unit commitment in electricity markets are presented in Bertsimas et al. (2012), Jiang et al. (2011), Jiang et al. (2014), Zhao and Zeng (2012), which study the problem under uncertain load or production from stochastic sources. A cutting-plane approach is used in all these papers to exploit the convex dependence of the objective function on the first-stage decision. Then, the max-min problem, resulting from the sequential enforcement of the worst-case parameter realization and decision on the optimal recourse action, is solved either employing an heuristic approach (see Bertsimas et al. (2012)) or as a Mixed Integer Linear Programming (MILP) problem after linearization of bilinear terms (see Jiang et al. (2011); Zhao and Zeng (2012)). Instead, Jiang et al. (2014) implement both approaches to compare their efficiency.

Jabr (2013), Ruiz and Conejo (2015), Mínguez and García-Bertrand (2016), and Zhang and Conejo (2018) apply ARO to the transmission expansion planning problem. Jabr (2013) models transmission network expansion taking into account uncertain renewable generation and load. The ARO problem is solved using a Benders decomposition scheme, which implies the computation of dual variables. On the contrary, Ruiz and Conejo (2015) apply a column-and-constraint generation method, based on primal cuts, to investigate how transmission expansion decisions change when demand growth, stochastic generation levels, and equipment failures are uncertain. In both papers, the ARO

problem is solved by decomposing it and iteratively solving a subproblem and a master problem. Mínguez and García-Bertrand (2016) propose a combination of the solution approaches used by Jabr (2013) and Ruiz and Conejo (2015). In particular, they account for a slightly modified version of the subproblem proposed by Jabr (2013) that is combined with the master problem formulated by Ruiz and Conejo (2015). The uncertain parameters considered are demand levels, capacity, and availability of generation facilities. Zhang and Conejo (2018) adopt an ARO approach to investigate the transmission expansion planning problem where both long-term uncertainty in demand growth and future production capacity, and short-term variability in demand and renewable power production levels are considered. The proposed ARO model is solved using a column-and-constraint generation method.

Robust optimization has been also used in other energy-related problems. For example, Baringo and Conejo (2010), Dominguez et al. (2012), and Baringo and Baringo (2016) propose robust models to derive optimal market offering strategies for power producers. In particular, Baringo and Conejo (2010) provide an approach to build hourly offering curves for a price-taker producer participating in a pool. Dominguez et al. (2012) focus their attention on a technique to build offering curves for a concentrating solar power plant. On the other side, Baringo and Baringo (2016) develop an ARO approach for the offering strategy of a virtual power plant composed of a conventional power plant, a wind-power unit, a storage facility, and flexible demands, participating in the day-ahead and the real-time markets as a single entity in order to optimize the use of their energy resources. Moreover, Baringo and Baringo (2017) use an ARO to deal with the generation and transmission expansion planning problem. Other applications of ARO problems include energy and reserve dispatch in electricity markets (see Moreira et al. (2015); Zugno and Conejo (2015)), the self-scheduling and trading for a virtual power plant (see Lima et al. (2015)), the integration of large-scale renewable power plants in energy systems (see Peng et al. (2014)), and the commitment and dispatch of heat and power units (see Zugno et al. (2016)).

Regarding expansion planning problems with stochastic renewable energy resources it is worthy mentioning the works by Khaligh et al. (2018) and Khaligh and Anvari-Moghaddam (2019), which propose a multi-attribute decision-making method and a stochastic decentralized-based approach, respectively, for the co-expansion planning of gas and electricity systems, as well as the work by Navidi et al. (2019), which proposes a bi-level fuzzy multi-criteria optimization method for the sub-transmission network expansion planning problem considering regional energy systems. Despite of their relevance, no one of these references deals with the expansion planning problem of a distribution system.

Considering models in power distribution systems, Ebrahimi and Amjady (2019) provide an ARO framework to address the day-ahead scheduling of a microgrid with wind, solar, batteries, and micro-turbine distributed generators, taking into account the associated uncertainties and AC power flow constraints. Regarding the expansion planning of distribution systems, there is an extensive literature. Contreras et al. (2016) provide an overview of different algorithms for joint expansion planning of distributed generation and distribution network models. Amjady et al. (2017) describe an ARO model for the expansion planning of a distribution systems; however, neither EVs nor storage units are considered. Asensio et al. (2016) propose a stochastic model for the joint distribution and generation expansion planning for isolated systems considering demand response and energy storage. de Quevedo et al. (2017) provide a stochastic multistage distribution expansion planning model that considers investments in distribution network assets, renewable units, energy storage, and EV charging stations. Najafi et al. (2019) propose a stochastic model to improve the resilience of joint power and water distribution networks. Munoz-Delgado et al. (2019) propose a three-level problem for the dynamic investment decision-making problem arising in a distribution system within a transactive energy environment.

Note that, despite the relevance of the above works, there are no references in the ARO literature that deal with EVs in a distribution system where investments in PV power plants, storage units, and charging stations are allowed. de Quevedo et al. (2017) consider also the investment in all these units; however, they propose a stochastic model that disregards long-term uncertainties. On the other hand, as in this paper, Amjady et al. (2017) propose an ARO model but neither EVs nor short-term variability are considered. Thus, there is a research gap that this chapter tries to fill.

Under this context, the contributions of this work are fourfold:

1. To develop an ARO model describing investments in a distribution system, where both short-term variability and long-term uncertainty are considered.
2. To model investments in PV power plants, storage units, and charging stations in an integrated framework.
3. To conduct a set of sensitivity analyses to evaluate how investment decisions vary on the basis of different levels of long-term uncertainty, the revenues of the central planner accruing from selling electricity to EVs at charging stations, the degree of autonomy of the considered distribution system from the main grid, and the central planner's investment budget availability.

4. To quantify the amount of CO₂ saved with the utilization of EVs to evaluate whether the environment can benefit from the progressive decarbonization of the road-transport.
5. To compare the results of the proposed approach with those of a stochastic model in terms of computational performance.

The rest of the chapter is organized as follows. Section 3.2 describes the formulation of the expansion planning problem applied to a distribution system, starting from a deterministic approach and then transforming it into an ARO problem. Section 3.3 illustrates the method used to solve the ARO model, while Section 3.4 describes the case study and the considered input data. Section 3.5 is devoted to the analysis of the obtained results. Section 3.6 concludes the chapter with some final remarks. Finally, 4.3 reports the data used in the case study.

3.2 Problem Formulation

This section describes the formulation of the proposed expansion planning model applied to a distribution system. For the sake of clarity, firstly a deterministic instance of the problem is provided, which only accounts for short-term variability, while the future peak demand, the price of the electricity interchanged with the main grid, and the number of EVs are considered as known parameters. The long-term uncertainty of these parameters is then characterized and the problem is formulated using an ARO model.

3.2.1 Notation

The indexes, the sets, the parameters, and the variables used in the models are listed below. Note that if symbol (ω) is included after any of the parameters and variables below, it denotes the value of these parameters and variables in scenario ω .

Indexes

c	Charging stations.
d	Representative days (operating conditions).
l	Demands.
ℓ	Lines.

n	Nodes.
r	Renewable (PV) units.
s	Storage units.
t	Hours.
v	Iterations.
ω	Scenarios.

Sets

$r(\ell)/s(\ell)$	Receiving/sending-end node of the ℓ th line.
Ψ_n^C	Charging stations located at node n .
Ψ_n^G	Connection to the main grid located at node n .
Ψ_n^R	Renewable units located at node n .
Ψ_n^S	Storage units located at node n .

Parameters

C_{dt}^E	Cost of the power used to charge EVs in hour t of representative day d [\$/kW].
$\tilde{C}_{dt}^G$	Forecast value of the price of power exchanged with the main grid in hour t of representative day d [\$/kW].
$\hat{C}_{dt}^G$	Maximum increase of the price of power exchanged with the main grid in hour t of representative day d [\$/kW].
C_{ldt}^{LS}	Load-shedding cost of demand l in hour t of representative day d [\$/kW].
C_c^{IC}	Annualized investment cost in charging station c [\$].
C_r^{IR}	Annualized investment cost in renewable unit r [\$].
C_s^{IS}	Annualized investment cost in storage unit s [\$].
$\bar{C}^I$	Annualized investment budget [\$].
$\underline{E}_s^S/\bar{E}_s^S$	Minimum/Maximum energy that can be stored in storage unit s [kWh].
$\underline{P}^G/\bar{P}^G$	Minimum/Maximum active power that can be exchanged with the main grid [kW].
$\underline{Q}^G/\bar{Q}^G$	Minimum/Maximum reactive power that can be exchanged with the main grid [kvar].
K_{ldt}^D	Demand factor of load l in hour t of representative day d [pu].

K_{cdt}^{EV}	EV availability factor in charging station c in hour t of representative day d [pu].
$K_{r dt}^R$	Capacity factor of renewable unit r in hour t of representative day d [pu].
N_d^D	Number of days grouped in representative day d [days].
$\tilde{N}_{cd}^{EV}$	Forecast number of electric vehicles in charging station c in representative day d .
$\hat{N}_{cd}^{EV}$	Maximum increase in the number of EVs in charging station c in representative day d .
$\overline{P}_c^C$	Charging capacity of charging station c .
$\tilde{P}_l^D$	Forecast peak load of demand l [kW].
$\hat{P}_{ldt}^D$	Maximum increase of the peak load of demand l [kW].
P_{cd}^{EV}	Maximum demand in charging station c in representative day d [kW].
P_{cd}^{EVmax}	Maximum daily demand in charging station c in representative day d [kW].
$\overline{P}_r^R$	Capacity of renewable unit r [pu].
$\overline{P}_s^{SC}$	Charging capacity of storage unit s [kW].
$\overline{P}_s^{SD}$	Discharging capacity of storage unit s [kW].
$\underline{V}_n/\overline{V}_n$	Minimum /Maximum voltages at node n [p.u.].
V_0	Substation voltage [p.u.].
α_{ldt}	ratio of reactive load of demand l in hour t of representative day d .
χ_ℓ	inductance of line ℓ .
κ_ℓ	resistance of line ℓ .
η_s^{SC}	Charging efficiency of storage unit s [%].
η_s^{SD}	Discharging efficiency of storage unit s [%].
Δt	Duration of time periods [h].
Λ^G	Uncertainty budget for the price of the power exchanged with the main grid.
Λ^L	Uncertainty budget for the peak demand.
Λ^{EV}	Uncertainty budget for the number of EVs.

Uncertain Variables

$\overline{c}_{dt}^G$	Price of the power exchanged with the main grid in hour t of representative day d [\$/kW].
$\overline{p}_l^D$	Peak load of demand l [kW].
$\overline{n}_{cd}^{EV}$	Number of EV in charging station c in representative day d .

Optimization Variables

e_{sdt}	Energy stored in storage unit s in hour t of representative day d [kWh].
p_{cdt}^C	Charging power in charging station c in hour t of representative day d [kW].
p_{gdt}^G	Active power exchanged with the main grid g in hour t of representative day d [kW].
q_{gdt}^G	Reactive power exchanged with the main grid g in hour t of representative day d [kvar].
$p_{\ell dt}^L$	Active flows through line ℓ in hour t of representative day d [kW; kvar].
$q_{\ell dt}^L$	Reactive flows through line ℓ in hour t of representative day d [kW; kvar].
p_{ldt}^{LS}	Load shed by demand l in hour t of representative day d [kW].
p_{rdt}^R	Active power from renewable unit r in hour t of representative day d [kW].
q_{rdt}^R	Reactive power from renewable unit r in hour t of representative day d [kW].
p_{sdt}^{SC}	Charging power of storage unit s in hour t of representative day d [kW].
p_{sdt}^{SD}	Discharging power of storage unit s in hour t of representative day d [kW].
u^{G+}	Binary variable that it is equal to 1 if the worst-case power price from the main grid is equal to its upper bound, otherwise it is equal to 0.
u^{G-}	Binary variable that it is equal to 1 if the worst-case power price from the main grid is equal to its lower bound, otherwise it is equal to 0.
u_l^{D+}	Binary variable that it is equal to 1 if the worst-case consumption profile of peak demand l is equal to its upper bound, otherwise it is equal to 0.
u_l^{D-}	Binary variable that it is equal to 1 if the worst-case consumption profile of peak demand l is equal to its lower bound, otherwise it is equal to 0.
u_c^{EV+}	Binary variable that it is equal to 1 if the worst-case utilization of charging stations c by EVs is equal to its upper bound, otherwise it is equal to 0.
u_c^{EV-}	Binary variable that it is equal to 1 if the worst-case utilization of charging stations c by EVs is equal to its lower bound, otherwise it is equal to 0.
x_c^C	Integer variable representing the number of charging stations c that are built.
x_r^R	Integer variable representing the number of renewable units r that are built.
x_s^S	Integer variable representing the number of storage units s that are built.
v_{ndt}	Voltage magnitude at node n in hour t of representative day d [p.u.].
θ	Auxiliary variable of the master problem.

3.2.2 Planning horizon and short-term variability characterization

The distribution expansion planning problem is solved for a long-term planning horizon, e.g., 30 years. However, as is customary in the literature (e.g., see Zhang and Conejo (2017)) to characterize the variability in the demand, production of PV units, price of electricity exchanged with the main grid, and EV behavior, a set of representative days are used to represent a future single target year. Each of these representative days defines the daily variability of the demand, the production of the stochastic PV units, the price of the electricity exchanged with the main grid, and the EV driving requirements.

Operating conditions used in this chapter have been defined using historical data and a modified version of the K-means clustering technique proposed in Baringo and Conejo (2013a). This method groups together historical data that are similar while maintaining the information as well as the auto- and cross-correlation among them. This method allows generating representative days by following the steps below:

1. Consider historical data of demand, renewable production, electricity prices, and EV driving requirements.
2. Select the number of representative days according to the needs of the problem. A low number of representative days may originate a poor representation of the historical data while a high number may lead to an unacceptable computational performance.
3. Randomly define the initial centroid of each representative day.
4. Compute the quadratic distance between each original historical day and the centroid of each representative day.
5. Allocate each historical day to the closest representative day according to the distances calculated in step 4.
6. Recalculate the centroid of each representative day using the historical days allocated to each representative day.

Then, steps 4-6 are repeated iteratively until there are no changes in the composition of representative days between two consecutive iterations.

The output of the K-means clustering technique is a set of representative days indexed by d , each one defined by 24 operating conditions (demand, renewable production, electricity prices, and EV

requirements), one for each hour of the day indexed by t , as well as the number of historical days allocated to each representative day (N_d^D), which is used as a measure of the weight of each representative day in the problem. Further information can be found in García-Cerezo et al. (2020).

Note that the output of this K-means method is used as input data of the planning model described in the following section. However, alternative methods such as the one developed by Pineda and Morales (2018) may be considered.

3.2.3 Deterministic model

The expansion planning problem considering a deterministic approach is formulated in compact form as the following MILP model:

$$\min_{\mathbf{x}, \mathbf{y}} \mathbf{C}^T \mathbf{x} + \mathbf{D}^T \mathbf{y} \quad (3.1a)$$

subject to

$$\mathbf{x} \in \mathbb{Z}^n, \quad (3.1b)$$

$$\mathbf{C}^T \mathbf{x} \leq \mathbf{E}, \quad (3.1c)$$

$$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y} = \mathbf{F}, \quad (3.1d)$$

$$\mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \leq \mathbf{I}, \quad (3.1e)$$

where $\mathbf{x}$ and $\mathbf{y}$ are vectors of variables representing expansion and operating decisions, respectively, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{I}$ are vectors and matrices with constant parameters dependent on the system characteristics, and n is the number of expansion options.

Problem (3.1) is driven by the minimization of both investment and operating costs in (3.1a). Constraints (3.1b) and (3.1c) are investment constraints, while (3.1d) and (3.1e) are operating constraints. A detailed formulation of this deterministic problem is provided in below.

$$\begin{aligned} \min_{\Phi^D} & \sum_s C_s^{\text{IS}} x_s^{\text{S}} + \sum_r C_r^{\text{IR}} x_r^{\text{R}} + \sum_c C_c^{\text{IC}} x_c^{\text{C}} \\ & + \sum_d N_d^D \left[\sum_t \left(\sum_g \tilde{C}_{dt}^{\text{G}} p_{gdt}^{\text{G}} + \sum_l C_l^{\text{LS}} p_{ldt}^{\text{LS}} - \sum_c C_{dt}^{\text{E}} p_{cdt}^{\text{C}} \right) \right] \end{aligned} \quad (3.2a)$$

subject to

$$x_s^S \in \mathbb{Z}^+, \quad \forall s, \quad (3.2b)$$

$$x_r^R \in \mathbb{Z}^+, \quad \forall r, \quad (3.2c)$$

$$x_c^C \in \mathbb{Z}^+, \quad \forall c, \quad (3.2d)$$

$$\sum_s C_s^{\text{IS}} x_s^S + \sum_r C_r^{\text{IR}} x_r^R + \sum_c C_c^{\text{IC}} x_c^C \leq \bar{C}^{\text{I}}, \quad (3.2e)$$

$$\begin{aligned} & \sum_{g \in \Psi_n^G} p_{gdt}^G + \sum_{r \in \Psi_n^R} p_{rdt}^R + \sum_{s \in \Psi_n^S} p_{sdt}^{\text{SD}} \eta_s^{\text{SD}} - \sum_{\ell | s(\ell)=n} p_{\ell dt}^L + \sum_{\ell | r(\ell)=n} p_{\ell dt}^L \\ &= \sum_{l \in \Psi_n^L} \left(\tilde{P}_l^{\text{D}} K_{ldt}^{\text{D}} - p_{ldt}^{\text{LS}} \right) + \sum_{s \in \Psi_n^S} p_{sdt}^{\text{SC}} + \sum_{c \in \Psi_n^C} p_{cdt}^C, \quad \forall n, \forall d, \forall t, \end{aligned} \quad (3.2f)$$

$$\sum_{g \in \Psi_n^G} q_{gdt}^G + \sum_{r \in \Psi_n^R} q_{rdt}^R - \sum_{\ell | s(\ell)=n} q_{\ell dt}^L + \sum_{\ell | r(\ell)=n} q_{\ell dt}^L = \sum_{l \in \Psi_n^L} \alpha_{ldt} \left(\tilde{P}_l^{\text{D}} K_{ldt}^{\text{D}} - p_{ldt}^{\text{LS}} \right), \quad \forall n, \forall d, \forall t, \quad (3.2g)$$

$$\underline{V}_n \leq v_{ndt} \leq \bar{V}_n, \quad \forall n/n : \text{ref.}, \forall d, \forall t, \quad (3.2h)$$

$$v_{ndt} = V_0, \quad n : \text{ref.}, \forall d, \forall t, \quad (3.2i)$$

$$v_{r(\ell)dt} = v_{s(\ell)dt} - \frac{\kappa_\ell p_{\ell dt}^L + \chi_\ell q_{\ell dt}^L}{V_0} \quad \forall \ell, \forall d, \forall t, \quad (3.2j)$$

$$\underline{P}^G \leq p_{gdt}^G \leq \bar{P}^G, \quad \forall g, \forall d, \forall t \quad (3.2k)$$

$$\underline{Q}^G \leq q_{gdt}^G \leq \bar{Q}^G, \quad \forall g, \forall d, \forall t \quad (3.2l)$$

$$0 \leq p_{ldt}^{\text{LS}} \leq \tilde{P}_l^{\text{D}} K_{ldt}^{\text{D}} \quad \forall l, \forall d, \forall t, \quad (3.2m)$$

$$0 \leq p_{rdt}^R \leq K_{rdt}^R \bar{P}_r^R x_r^R, \quad \forall r, \forall d, \forall t, \quad (3.2n)$$

$$e_{sdt} = e_{sd,t-1} + p_{sdt}^{\text{SC}} \eta_s^{\text{SC}} \Delta t - p_{sdt}^{\text{SD}} \Delta t, \quad \forall s, \forall d, \forall t \quad (3.2o)$$

$$\underline{E}_s^S x_s^S \leq e_{sdt} \leq \bar{E}_s^S x_s^S, \quad \forall s, \forall d, \forall t, \quad (3.2p)$$

$$0 \leq p_{sdt}^{\text{SC}} \leq \bar{P}_s^{\text{SC}} x_s^S, \quad \forall s, \forall d, \forall t, \quad (3.2q)$$

$$0 \leq p_{sdt}^{\text{SD}} \leq \bar{P}_s^{\text{SD}} x_s^S, \quad \forall s, \forall d, \forall t, \quad (3.2r)$$

$$0 \leq p_{cdt}^C \leq \bar{P}_c^C x_c^C, \quad \forall c, \forall d, \forall t, \quad (3.2s)$$

$$0 \leq p_{cdt}^C \leq \tilde{N}_{cd}^{\text{EV}} P_{cd}^{\text{EV}} K_{cdt}^{\text{EV}}, \quad \forall c, \forall d, \forall t, \quad (3.2t)$$

$$\sum_t p_{cdt}^C \leq \tilde{N}_{cd}^{\text{EV}} P_{cd}^{\text{EVmax}}, \quad \forall c, \forall d, \quad (3.2u)$$

where set $\Phi^D = \{x_s^S, \forall s; x_r^R, \forall r; x_c^C, \forall c; p_{gdt}^G, \forall d, \forall t; q_{gdt}^G, \forall d, \forall t; p_{rdt}^R, \forall r, \forall d, \forall t; q_{rdt}^R, \forall r, \forall d, \forall t; p_{sdt}^{SD}, p_{sdt}^{SC}, e_{sdt}, \forall s, \forall d, \forall t; p_{ldt}^L, \forall l, \forall d, \forall t; q_{ldt}^L, \forall l, \forall d, \forall t; p_{ldt}^{LS}, \forall l, \forall d, \forall t; p_{cdt}^C, \forall c, \forall d, \forall t; v_{ndt}, \forall n, \forall d, \forall t\}$ includes the optimization variables of problem (3.2).

Problem (3.2) is driven by the minimization of both investment and operating costs. Objective function (3.18a) includes the following terms:

1. Term $\sum_s C_s^{IS} x_s^S + \sum_r C_r^{IR} x_r^R + \sum_c C_c^{IC} x_c^C$ represents the annualized investment costs in storage, renewable units, and charging stations.
2. Terms $\sum_t \left(\sum_g \tilde{C}_{dt}^G p_{gdt}^G + \sum_l C_l^{LS} p_{ldt}^{LS} \right), \forall d$, include the costs (revenues) of the power bought (sold) from the main grid and the load-shedding cost. No charges are foreseen for the solar PV power production. Note that p_{gdt}^G is a free variable: when it is positive, it indicates the amount of electricity that the distribution system withdraws from the main grid, if negative it stands for the amount of power that the distribution system sells back to the main grid.
3. Terms $\sum_t \sum_c C_{dt}^E p_{cdt}^C, \forall d$, represent the revenues achieved by selling energy to EVs through charging stations.

Note that terms in items 2 and 3 depend on representative days d and are multiplied by factor N_d^D to make investment and operating costs comparable.

Constraints of the problem can be classified into two groups. Constraints (3.18b)-(3.18e) are investment constraints, while (3.18f)-(3.18u) are operating constraints.

Constraints (3.18b), (3.18c), and (3.18d) are declarations of integer variables representing the investment in storage, PV, and charging units, respectively. Note that these variables identify the number of plants that can be built. We fix, “a priori”, the capacity of these units to reflect the fact that residential scale or commercial scale installation have industry standards. The investments costs indicated in the objective function are proportional to the capacity that the specific unit can assume. Constraint (3.18e) represents the investment budget. Constraints (3.18f) and (3.18g) impose the power balances at each node for active and reactive power, respectively. Conditions (3.18f) state that the amount of active power exchanged with main grid in addition to the power produced by PV installations, discharged from the storage units and adjusted by the flows with the other nodes has to be equal to the sum of consumers’ demand corrected by possible load shedding and the electricity required by charging storage units to supply EVs. Constraints (3.18g) have a similar meaning. Constraints (3.18h) and (3.18i) bound voltage magnitude and fix the voltage angle at

the reference node, respectively (see Ebrahimi and Amjady, 2019). The voltage relation between connected nodes is enforced by conditions (3.18j). On the other side, constraints (3.18k) and (3.18l) limit the active and reactive power that can be exchanged with the main grid, respectively. Note that considering a complete AC power model would result in a nonlinear optimization problem for which optimality is not guaranteed. However, after making the expansion plans, the distribution system operator should run a full AC power flow model to guarantee the feasible operation of the distribution system. Constraints (3.18m) and (3.18n) limit the load shed and the production of renewable units, respectively. In particular, the load shedding has to be lower than the consumers' demand that is computed as the product between the peak demand parameter $\tilde{P}_l^D$ and the load factor K_{ldt}^D computed through the operating conditions. Conditions (3.18n) have been constructed in a similar way. The term $K_{rdt}^R \bar{P}_r^R x_r^R$ identifies the maximum electricity production of these stochastic units. In particular, the solar availability factor K_{rdt}^R is multiplied by $\bar{P}_r^R x_r^R$ that indicates the total capacity of the PV units installed in a node of the distribution system. This is obtained by the product between the investment variable x_r^R and the assumed "a priori" capacity $\bar{P}_r^R$ that is attributed to each PV unit. Equations (3.18o)-(3.18r) model the working of storage units. Constraints (3.18o) define the energy evolution in storage units. Equations (3.18p) impose bounds on the energy stored. Constraints (3.18q) and (3.18r) limit the charging and discharging powers, respectively. The bounds of constraints (3.18p)-(3.18r) are obtained by multiplying the investment variables x_s^S , which express the number of storage stations to be built, by ad hoc parameters. Finally, equations (3.18s)-(3.18u) model the charging stations. Constraints (3.18s), (3.18t), and (3.18u) impose limits on the charging power in charging stations depending on the charging capacity, EV hourly demand, and EV daily demand, respectively. The bounds of constraints (3.18s) are computed as the product between the number of charging stations constructed x_c^C and the "a priori" capacity assigned to each charging unit identified by the parameter $\bar{P}_c^C$. Constraints (3.18t) and (3.18u) impose limits on the charging power in charging stations depending on the EV hourly demand and EV daily demand, respectively.

3.2.4 Long-term uncertainty characterization: uncertainty set

In the deterministic problem (3.1), there is a number of parameters, namely, the peak demand, the number of EVs, and the price of the power exchanged with the main grid that are assumed to be known at the time of making the expansion decisions. However, when these decisions are made, this information is not known by the expansion planner. Therefore, it is also necessary to represent

the long-term uncertainty in the decision making problem in order to obtain informed expansion decisions.

For the problem considered, uncertainty in the peak demand, the price of the power exchanged with the main grid, and the number of EVs is modeled by decision variables that take values within known confidence bounds:

$$\mathbf{v} \in \{\tilde{\mathbf{v}} - \hat{\mathbf{v}}, \tilde{\mathbf{v}} + \hat{\mathbf{v}}\}, \quad (3.3)$$

where $\mathbf{v}$ is a vector representing the uncertain variables, while $\tilde{\mathbf{v}}$ and $\hat{\mathbf{v}}$ are respectively the vectors of forecast and fluctuation levels of these uncertain variables.

In the proposed ARO model, the worst-case uncertainty realization given the expansion decisions is identified, i.e., that uncertainty realization that maximizes the operation cost. For the polyhedral uncertainty set considered in this chapter and represented by equations (3.3), the worst-case uncertainty realization corresponds to a vertex of the polyhedron representing the uncertainty set (see Jiang et al. (2014); Morales-Napoles (2010)). Thus, it is possible to use the following equivalent binary-variable-based set:

$$\Omega = \left\{ \mathbf{v} = \tilde{\mathbf{v}} + \text{diag}(\mathbf{u}^+) \hat{\mathbf{v}} - \text{diag}(\mathbf{u}^-) \hat{\mathbf{v}}, \right. \quad (3.4a)$$

$$\mathbf{u}^+, \mathbf{u}^- \in \{0, 1\}^m, \quad (3.4b)$$

$$\sum_{k=1}^m (u_k^+ + u_k^-) \leq \Lambda, \quad (3.4c)$$

$$u_k^+ + u_k^- \leq 1, \quad \forall k \}, \quad (3.4d)$$

where m indicates the size of the vector of variables $\mathbf{v}$.

Constraints (3.4a) express the uncertain variables in vector $\mathbf{v}$ in terms of the forecast and fluctuation levels. Constraints (3.4b) define the vectors of binary variables $\mathbf{u}^+$ and $\mathbf{u}^-$. Constraints (3.4c) allow controlling the robustness in the solution through the so-called uncertainty budget Λ . If this uncertainty budget is equal to 0, it means that all uncertain variables in vector $\mathbf{v}$ are equal to their forecast values, i.e., uncertainty is disregarded. As the value of this uncertainty budget is increased, the uncertain variables in vector $\mathbf{v}$ are allowed to deviate from their forecast values, i.e., a comparatively more robust solution is considered. Finally, constraints (3.4d) impose that each uncertain variable in vector $\mathbf{v}$ cannot be simultaneously equal to its lower and upper bounds. A detailed formulation of this uncertainty set is provided below. Uncertainty in the peak demand, the

price of the power exchanged with the main grid, and the number of EVs is modeled by decision variables that take values within known confidence bounds:

$$\bar{c}_{dt}^G \in \{\tilde{C}_{dt}^G - \hat{C}_{dt}, \tilde{C}_{dt}^G + \hat{C}_{dt}^G\}, \quad \forall d, \forall t, \quad (3.5a)$$

$$\bar{p}_l^D = \{\tilde{P}_l^D - \hat{P}_l^D, \tilde{P}_l^D + \hat{P}_l^D\}, \quad \forall l, \quad (3.5b)$$

$$\bar{n}_{cd}^{EV} = \{\tilde{N}_{cd}^{EV} - \hat{N}_{cd}^{EV}, \tilde{N}_{cd}^{EV} + \hat{N}_{cd}^{EV}\}, \quad \forall c, \forall d. \quad (3.5c)$$

As explained in Section 3.2.4, it is possible to use the following equivalent binary-variable-based uncertainty set Ω :

$$\Omega = \{\bar{c}_{dt}^G = \tilde{C}_{dt}^G + u^{G+} \hat{C}_{dt}^G - u^{G-} \hat{C}_{dt}^G, \quad \forall d, \forall t, \quad (3.6a)$$

$$\bar{p}_l^D = \tilde{P}_l^D + u_l^{D+} \hat{P}_l^D - u_l^{D-} \hat{P}_l^D, \quad \forall l, \quad (3.6b)$$

$$\bar{n}_{cd}^{EV} = \tilde{N}_{cd}^{EV} + u_c^{EV+} \hat{N}_{cd}^{EV} - u_c^{EV-} \hat{N}_{cd}^{EV}, \quad \forall c, \forall d, \quad (3.6c)$$

$$u^{G+}, u^{G-} \in \{0, 1\}, \quad (3.6d)$$

$$u_l^{D+}, u_l^{D-} \in \{0, 1\}, \quad \forall l, \quad (3.6e)$$

$$u_c^{EV+}, u_c^{EV-} \in \{0, 1\}, \quad \forall c, \quad (3.6f)$$

$$u^{G+} + u^{G-} \leq 1, \quad (3.6g)$$

$$u_l^{D+} + u_l^{D-} \leq 1, \quad \forall l, \quad (3.6h)$$

$$u_c^{EV+} + u_c^{EV-} \leq 1, \quad \forall c, \quad (3.6i)$$

$$u^{G+} + u^{G-} \leq \Lambda^G, \quad (3.6j)$$

$$\sum_l (u_l^{D+} + u_l^{D-}) \leq \Lambda^D, \quad (3.6k)$$

$$\sum_c (u_c^{EV+} + u_c^{EV-}) \leq \Lambda^{EV}. \quad (3.6l)$$

Constraints (3.6a), (3.6b), and (3.6c) express the price of the power bought from the main grid, the peak demand, and the number of EVs in terms of the forecast and fluctuation levels of the corresponding uncertain variables, respectively. Constraints (3.6d), (3.6e), and (3.6f) define binary variables. Constraints (3.6g), (3.6h), and (3.6i) impose that uncertain variables cannot be simultaneously at both their upper and lower bounds. Finally, constraints (3.6j), (3.6k), and (3.6l) allow controlling the robustness in the solution through the so-called uncertainty budgets Λ^G , Λ^D , and Λ^{EV} .

Note that the above uncertainty set neglects the possible correlation among long-term uncertainties. In this sense, the use of ellipsoidal uncertainty sets as Golestaneh et al. (2018) or multivariate prediction polyhedral as Golestaneh et al. (2018) constitute promising strategies. This is subject of future research.

3.2.5 Adaptive robust optimization model

The proposed ARO model is formulated as a three-level programming problem. The first level determines the expansion decisions minimizing both the expansion and the operation costs. Given these expansion decisions, the second level identifies the worst-case realization of peak demands, number of EVs, and price of the electricity exchanged with the main grid that maximizes the operation costs. Finally, the third level models the operation of the distribution system minimizing the operation cost for given first- and second-level decisions. The proposed three-level programming problem is formulated in compact form as follows:

$$\min_{\mathbf{x}} \mathbf{C}^T \mathbf{x} + \max_{\mathbf{v} \in \Omega} \min_{\mathbf{y} \in \Xi(\cdot)} \mathbf{D}(\mathbf{v})^T \mathbf{y} \quad (3.7a)$$

subject to

$$\text{Constraints (3.1b) -- (3.1c).} \quad (3.7b)$$

Problem (3.7) involves three nested optimization problems:

1. The first level associated with the expansion decisions, i.e., variables in vector $\mathbf{x}$.
2. The second level related to the worst-case realization of the price of the electricity exchanged with the main grid, the peak demand, and the number of EVs, i.e., variables in vector $\mathbf{v}$.
3. The third level modeling the reaction of the distribution system against first- and second-level decisions, i.e., variables in vector $\mathbf{y}$.

Problem (3.7) is driven by the minimization of the worst-case expansion and operation costs (3.7a) subject to the expansion constraints (3.7b) described in Section 3.2.3. In problem (3.7), Ω and Ξ are the uncertainty and the feasibility sets, respectively. Uncertainty set Ω is described in Section

3.2.4, while set Ξ identifies the feasible space of the third-level optimization variables as explained in Section 3.2.6. The proposed ARO problem is formulated as follows:

$$\begin{aligned} & \min_{\Phi^{L1}} \sum_s C_s^{IS} x_s^S + \sum_r C_r^{IR} x_r^R + \sum_c C_c^{IC} x_c^C \\ & + \max_{\Phi^{L2} \in \Omega} \min_{\Phi^{L3} \in \Xi(\cdot)} \\ & \sum_d N_d^D \left[\sum_t \left(\sum_g \bar{c}_{dt}^G p_{gdt}^G + \sum_l C_l^{LS} p_{ldt}^{LS} - \sum_c C_{dt}^{EC} p_{cdt}^C \right) \right] \end{aligned} \quad (3.8a)$$

subject to

$$\text{Constraints (3.18b) – (3.18e)}. \quad (3.8b)$$

Problem (3.8) involves three nested optimization problems:

1. The first level associated with the expansion decisions, i.e., variables in set $\Phi^{L1} = \{x_s^S, \forall s; x_r^R, \forall r; x_c^C, \forall c\}$.
2. The second level related to the worst-case realization of the value of the electricity exchanged with the main grid, the peak demand, the number of EVs, i.e., variables in set $\Phi^{L2} = \{\bar{c}_{dt}^G, \forall d, \forall t; \bar{p}_l^D, \forall l; \bar{n}_{cd}^{EV}, \forall c, \forall d\}$.
3. The third level modeling the reaction of the distribution system against first- and second-level decisions, i.e., variables in set $\Phi^{L3} = \{e_{sdt}^S, \forall s; p_{cdt}^C, \forall c; p_{gdt}^G; q_{gdt}^G; p_{ldt}^{LS}, \forall l; p_{ldt}^L, \forall l; q_{ldt}^L, \forall l; p_{rdt}^R, \forall r; q_{rdt}^R, \forall r; p_{sdt}^{SC}, \forall s; p_{sdt}^{SD}, \forall s; \delta_{ndt}, \forall n\}, \forall d, \forall t$.

As explained in Section 3.2.5, problem (3.8) is driven by the minimization of the worst-case expansion and operation costs (3.8a) subject to the expansion constraints (3.8b). In problem (3.8), Ω and Ξ are the uncertainty and the feasibility sets, respectively. Uncertainty set Ω is described in 3.2.4, while set Ξ identifies the feasible space of the third-level optimization variables as explained in 3.2.6.

3.2.6 Definition of feasibility sets

Given first- and second-level decision variables, set Ξ models the feasible space of third-level optimization variables:

$$\Xi(\mathbf{x}, \mathbf{v}) = \left\{ \mathbf{y} : \right.$$

$$\mathbf{Ax} + \mathbf{By} = \mathbf{F}(\mathbf{v}) : \boldsymbol{\lambda}, \quad (3.9a)$$

$$\mathbf{Gx} + \mathbf{Hy} \leq \mathbf{I}(\mathbf{v}) : \boldsymbol{\mu}, \quad (3.9b)$$

$$\left. \right\}.$$

Note that the feasibility set Ξ is parameterized in terms of first- and second-level decision variables. Constraints (3.9a)-(3.9b) are identical to constraints (3.1c)-(3.1d). Finally, observe that the dual variables are provided following a colon. A detailed formulation of this feasibility set is provided below. Given first- and second-level decision variables, set Ξ models the feasible space of third-level optimization variables:

$$\Xi \left(x_s^S, \forall s; x_r^R, \forall r; x_c^C, \forall c; \bar{c}_{dt}^G, \forall d, \forall t; \bar{p}_{ldt}^D, \forall l, \forall d, \forall t; \bar{n}_{cd}^{EV}, \forall c, \forall d \right) = \{ \Phi_{dt}^{L3} :$$

$$\begin{aligned} & \sum_{g \in \Psi_n^G} p_{gdt}^G + \sum_{r \in \Psi_n^R} p_{rdt}^R + \sum_{s \in \Psi_n^S} p_{sdt}^{SD} \eta^{SD} - \sum_{\ell | s(\ell)=n} p_{ldt}^L + \sum_{\ell | r(\ell)=n} p_{ldt}^L \\ & = \sum_{l \in \Psi_n^L} \left(K_{ldt}^D \bar{p}_{ldt}^D - p_{ldt}^{LS} \right) + \sum_{s \in \Psi_n^S} p_{sdt}^{SC} + \sum_{c \in \Psi_n^C} p_{cdt}^C : \lambda_{ndt}^P, \quad \forall n, \forall d, \forall t, \end{aligned} \quad (3.10a)$$

$$\sum_{g \in \Psi_n^G} q_{gdt}^G + \sum_{r \in \Psi_n^R} q_{rdt}^R - \sum_{\ell | s(\ell)=n} q_{ldt}^L + \sum_{\ell | r(\ell)=n} q_{ldt}^L = \sum_{l \in \Psi_n^L} \alpha_{ldt} \left(\tilde{P}_l^D K_{ldt}^D - p_{ldt}^{LS} \right) : \lambda_{ndt}^Q, \quad \forall n, \forall d, \forall t, \quad (3.10b)$$

$$\underline{V}_n \leq v_{ndt} \leq \bar{V}_n : \underline{\mu}_{ndt}^N, \bar{\mu}_{ndt}^N, \quad \forall n/n : \text{ref.}, \forall d, \forall t, \quad (3.10c)$$

$$v_{ndt} = V_0, \quad n : \text{ref.}, \forall d, \forall t, \quad (3.10d)$$

$$v_{r(\ell)dt} = v_{s(\ell)dt} - \frac{\kappa_{\ell} p_{ldt}^L + \chi_{\ell} q_{ldt}^L}{V_0} : \varphi_{ldt}^L, \quad \forall \ell, \forall d, \forall t, \quad (3.10e)$$

$$\underline{P}^G \leq p_{gdt}^G \leq \bar{P}^G : \underline{\mu}_{gdt}^P, \bar{\mu}_{gdt}^P, \quad \forall g, \forall d, \forall t \quad (3.10f)$$

$$\underline{Q}^G \leq q_{gdt}^G \leq \bar{Q}^G : \underline{\mu}_{gdt}^Q, \bar{\mu}_{gdt}^Q, \quad \forall g, \forall d, \forall t \quad (3.10g)$$

$$0 \leq p_{ldt}^{LS} \leq \bar{p}_l^D k_{ldt}^D : \underline{\mu}_{ldt}^D, \bar{\mu}_{ldt}^D, \quad \forall l, \forall d, \forall t, \quad (3.10h)$$

$$0 \leq p_{rdt}^R \leq K_{rdt}^R \bar{P}_r^R x_r^R : \underline{\mu}_{rdt}^R, \bar{\mu}_{rdt}^R, \quad \forall r, \forall d, \forall t, \quad (3.10i)$$

$$e_{sdt} = e_{sdt,t-1} + p_{sdt}^{SC} \eta_s^{SC} \Delta t - p_{sdt}^{SD} \Delta t : \varphi_{sdt}^E, \quad \forall s, \forall d, \forall t, \quad (3.10j)$$

$$\underline{E}_s^S x_s^S \leq e_{sdt} \leq \bar{E}_s^S x_s^S : \underline{\mu}_{sdt}^E, \bar{\mu}_{sdt}^E, \quad \forall s, \forall d, \forall t, \quad (3.10k)$$

$$0 \leq p_{sdt}^{SC} \leq \bar{P}_s^{SC} x_s^S : \underline{\mu}_{sdt}^{SC}, \bar{\mu}_{sdt}^{SC}, \quad \forall s, \forall d, \forall t, \quad (3.10l)$$

$$0 \leq p_{sdt}^{SD} \leq \bar{P}_s^{SD} x_s^S : \underline{\mu}_{sdt}^{SD}, \bar{\mu}_{sdt}^{SD}, \quad \forall s, \forall d, \forall t, \quad (3.10m)$$

$$0 \leq p_{cdt}^C \leq x_c^C \bar{P}_c^C : \underline{\mu}_{cdt}^C, \bar{\mu}_{cdt}^C, \quad \forall c, \forall d, \forall t, \quad (3.10n)$$

$$0 \leq p_{cdt}^C \leq \bar{n}_{cd}^{EV} P_{cd}^{EV} K_{cdt}^{EV} : \underline{\mu}_{cdt}^{EV}, \bar{\mu}_{cdt}^{EV}, \quad \forall c, \forall d, \forall t, \quad (3.10o)$$

$$\sum_t p_{cdt}^C \leq \bar{n}_{cd}^{EV} P_{cd}^{EVmax} : \varphi_{cd}^{EV}, \quad \forall c, \forall d \quad (3.10p)$$

}.

Note that the feasibility set Ξ is parameterized in terms of first- and second-level decision variables. Constraints (3.10a)-(3.10p) are identical to constraints (3.18f)-(3.18u). Finally, observe that the dual variables are provided following a colon.

3.3 Solution Procedure

The ARO problem (3.7) is a three-level optimization problem that is solved using a column-and-constraint generation algorithm whose convergence to the optimal solution is guaranteed (see Zeng and Zhao (2013)). This algorithm is based on the iterative solution of a master problem and a subproblem as explained below:

1. First, solve a master problem that is mainly based on the first-level problem of (3.7). The master problem uses as input data the outcomes of the subproblem (variables in vector $\mathbf{v}$, representing the worst-case uncertainty realization variables).
2. Then, solve a subproblem that is mainly based on the second-and third-level problems of (3.7). This subproblem uses as input data the outcomes of the master problem (variables in vector $\mathbf{x}$, representing the expansion decisions).
3. The master problem and the subproblem are iteratively solved until convergence to the optimal solution.

Figure 3.1 clarifies the iterative solution of the master problem and the subproblem that are further described in the following sections.

3.3.1 Master problem

The master problem at iteration v is provided below:

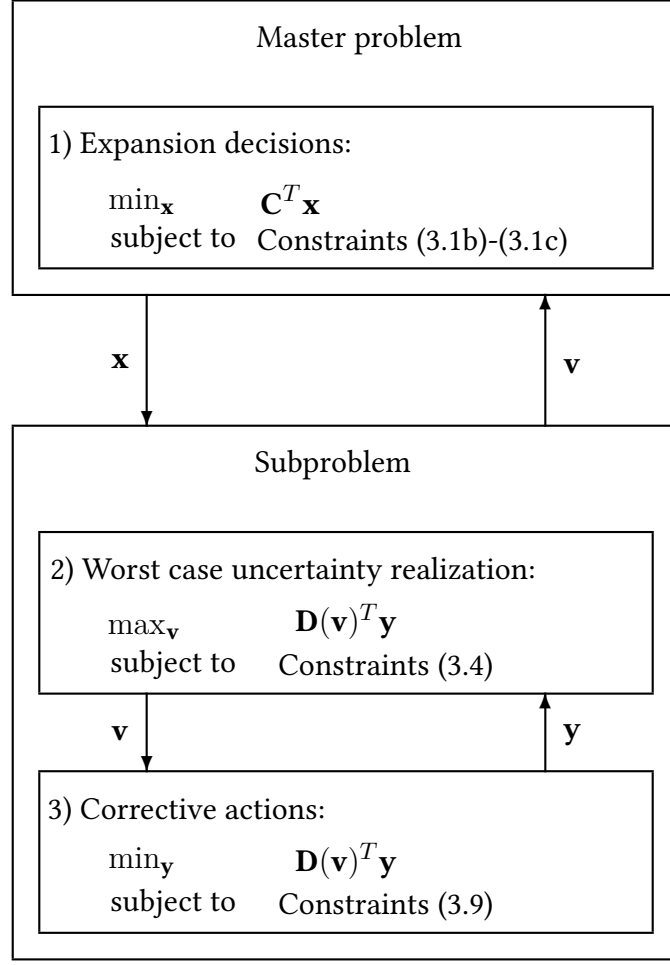


Figure 3.1: Iterative solution approach of the ARO problem

$$\min_{\Phi^M}$$

$$\mathbf{C}^T \mathbf{x} + \theta \tag{3.11a}$$

subject to

$$\text{Constraints (3.1b) -- (3.1c),} \tag{3.11b}$$

$$\theta \geq \mathbf{D}(\mathbf{v}^{(v')})^T \mathbf{y}_{v'}, \quad \forall v' \leq v, \tag{3.11c}$$

$$\mathbf{Ax} + \mathbf{By}_{v'} = \mathbf{F}(\mathbf{v}^{(v')}), \quad \forall v' \leq v, \tag{3.11d}$$

$$\mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y}_{v'} \leq \mathbf{I}(\mathbf{v}^{(v')}), \quad \forall v' \leq v, \quad (3.11e)$$

where variables in set $\Phi^M = \{\mathbf{x}, \theta, \mathbf{y}_{v'}, \forall v' \leq v\}$ are the optimization variables of master problem (3.11).

Master problem (3.11) is a relaxed version of the three-level problem (3.7) in which auxiliary variable θ iteratively approximates the worst-case value of the second-level objective function. Therefore, the size of master problem (3.11) increases with the number of iterations. Finally, note that parameters with superscript (v') refer to the optimal values of variables yielded by the subproblem at iteration v' . A detailed description of the master problem associated with the proposed ARO model is provided below.

The master problem at iteration v is provided below:

$$\min_{\Phi^M} \sum_s C_s^{\text{IS}} x_s^S + \sum_r C_r^{\text{IR}} x_r^R + \sum_c C_c^{\text{IC}} x_c^C + \sum_{\ell \in \Psi^{\text{L}+}} + \theta \quad (3.12a)$$

subject to

$$\text{Constraints (3.18b) -- (3.18e)} \quad (3.12b)$$

$$\theta \geq \sum_d N_d^{\text{D}} \left[\sum_t \left(\bar{c}_{dt}^{\text{G}(v')} \sum_g p_{gdtv'}^{\text{G}} + \sum_l C_l^{\text{LS}} p_{ldtv'}^{\text{LS}} - \sum_c C_{dt}^{\text{E}} p_{cdtv'}^{\text{C}} \right) \right], \quad \forall v' \leq v, \quad (3.12c)$$

$$\begin{aligned} & \sum_{g \in \Psi_n^{\text{G}}} p_{gdtv'}^{\text{G}} + \sum_{r \in \Psi_n^{\text{R}}} p_{rdtv'}^{\text{R}} + \sum_{s \in \Psi_n^{\text{S}}} p_{sdtv'}^{\text{SD}} \eta_s^{\text{SD}} - \sum_{\ell | s(\ell)=n} p_{ldtv'}^{\text{L}} + \sum_{\ell | r(\ell)=n} p_{ldtv'}^{\text{L}} \\ &= \sum_{l \in \Psi_n^{\text{L}}} \left(K_{ldt}^{\text{D}} \bar{p}_l^{\text{D}(v')} - p_{ldtv'}^{\text{LS}} \right) + \sum_{s \in \Psi_n^{\text{S}}} p_{sdtv'}^{\text{SC}} + \sum_{c \in \Psi_n^{\text{C}}} p_{cdtv'}^{\text{C}}, \\ & \forall n, \forall d, \forall t, \forall v' \leq v, \end{aligned} \quad (3.12d)$$

$$\begin{aligned} & \sum_{g \in \Psi_n^{\text{G}}} q_{gdtv'}^{\text{G}} + \sum_{r \in \Psi_n^{\text{R}}} q_{rdtv'}^{\text{R}} - \sum_{\ell | s(\ell)=n} q_{ldtv'}^{\text{L}} + \sum_{\ell | r(\ell)=n} q_{ldtv'}^{\text{L}} = \sum_{l \in \Psi_n^{\text{L}}} \alpha_{ldt} \left(\tilde{P}_l^{\text{D}} K_{ldt}^{\text{D}} - p_{ldtv'}^{\text{LS}} \right), \\ & \forall n, \forall d, \forall t, \forall v' \leq v, \end{aligned} \quad (3.12e)$$

$$\underline{V}_n \leq v_{ndtv'} \leq \bar{V}_n, \quad \forall n/n : \text{ref.}, \forall d, \forall t, \forall v' \leq v, \quad (3.12f)$$

$$v_{ndtv'} = V_0, \quad n : \text{ref.}, \forall d, \forall t, \forall v' \leq v, \quad (3.12g)$$

$$v_{r(\ell)dtv'} = v_{s(\ell)dtv'} - \frac{\kappa_{\ell} p_{ldtv'}^{\text{L}} + \chi_{\ell} q_{ldtv'}^{\text{L}}}{V_0} \quad \forall \ell, \forall d, \forall t, \forall v' \leq v, \quad (3.12h)$$

$$\underline{P}^G \leq p_{gdtv'}^G \leq \overline{P}^G, \quad \forall g, \forall d, \forall t, \forall v' \leq v, \quad (3.12i)$$

$$\underline{Q}^G \leq q_{gdtv'}^G \leq \overline{Q}^G, \quad \forall g, \forall d, \forall t, \forall v' \leq v, \quad (3.12j)$$

$$0 \leq p_{ldtv'}^{LS} \leq K_{ldt}^D \overline{P}_l^{D(v')}, \quad \forall l, \forall d, \forall t, \forall v' \leq v, \quad (3.12k)$$

$$0 \leq p_{rdtv'}^R \leq K_{rdt}^R \overline{P}_r^R x_r^R, \quad \forall r, \forall d, \forall t, \forall v' \leq v, \quad (3.12l)$$

$$e_{sdv'} = e_{sd,t-1,v'} + p_{sdv'}^{SC} \eta_s^{SC} \Delta t - p_{sdv'}^{SD} \Delta t, \quad \forall s, \forall d, \forall t, \forall v' \leq v, \quad (3.12m)$$

$$\underline{E}_s^S x_s^S \leq e_{sdv'} \leq \overline{E}_s^S x_s^S, \quad \forall s, \forall d, \forall t, \forall v' \leq v, \quad (3.12n)$$

$$0 \leq p_{sdv'}^{SC} \leq \overline{P}_s^{SC} x_s^S, \quad \forall s, \forall d, \forall t, \forall v' \leq v, \quad (3.12o)$$

$$0 \leq p_{sdv'}^{SD} \leq \overline{P}_s^{SD} x_s^S, \quad \forall s, \forall d, \forall t, \forall v' \leq v, \quad (3.12p)$$

$$0 \leq p_{cdv'}^C \leq x_c^C \overline{P}_c^C, \quad \forall c, \forall d, \forall t, \forall v' \leq v, \quad (3.12q)$$

$$0 \leq p_{cdv'}^C \leq \overline{n}_{cd}^{EV(v')} P_{cd}^{EV} K_{cd}^{EV}, \quad \forall c, \forall d, \forall t, \forall v' \leq v, \quad (3.12r)$$

$$\sum_t p_{cdv'}^C \leq \overline{n}_{cd}^{EV(v')} P_{cd}^{EVmax}, \quad \forall c, \forall d, \forall v' \leq v. \quad (3.12s)$$

where variables in set $\Phi^M = \{\Phi^{L1}; \theta; e_{sdv'}^S, \forall d, \forall t, \forall s, \forall v' \leq v; p_{cdv'}^C, \forall c, \forall d, \forall t, \forall v' \leq v; p_{gdtv'}^G, \forall g, \forall d, \forall t, \forall v' \leq v; q_{gdtv'}^G, \forall g, \forall d, \forall t, \forall v' \leq v; p_{ldtv'}^{LS}, \forall l, \forall d, \forall t, \forall v' \leq v; p_{ldtv'}^L, \forall l, \forall d, \forall t, \forall v' \leq v; q_{ldtv'}^L, \forall l, \forall d, \forall t, \forall v' \leq v; p_{rdtv'}^R, \forall r, \forall d, \forall t, \forall v' \leq v; q_{rdtv'}^R, \forall r, \forall d, \forall t, \forall v' \leq v; p_{sdv'}^{SC}, \forall s, \forall d, \forall t, \forall v' \leq v; p_{sdv'}^{SD}, \forall s, \forall d, \forall t, \forall v' \leq v; \delta_{ndtv'}, \forall n, \forall d, \forall t, \forall v' \leq v\}$ are the optimization variables of master problem (3.12).

3.3.2 Subproblem

At each iteration v , master problem (3.11) provides the expansion decisions $\mathbf{x}$. Given these decisions, the subproblem formulated below determines the worst-case uncertainty realization:

$$\max_{\mathbf{v} \in \Omega} \min_{\mathbf{y} \in \Xi(\cdot)} \mathbf{D}(\mathbf{v})^T \mathbf{y}. \quad (3.13)$$

Subproblem (3.13) is a bilevel model whose lower-level problem is continuous and linear on its decision variables. Thus, the lower-level problem can be replaced by its dual constraints. Moreover, using the strong duality equality, the objective function of the subproblem can be replaced by the dual lower-level objective function (Bertsimas et al. (2012)), which also entails the utilization of big-M constants. These big-M values are selected using a trial-and-error tuning procedure for the

sake of simplicity. As shown in Pineda and Morales (2019), this procedure is commonly used in the literature; however, it may lead to suboptimal solutions and high computational times in some cases. This issue will be subject of future research. As a result, the subproblem is finally formulated as a single-level MILP problem. At each iteration v , master problem (3.12) provides the expansion decisions. Given these decisions, the subproblem formulated below determines the worst-case uncertainty realization:

$$\max_{\Phi^{L2} \in \Omega} \min_{\Phi_{dt}^{L3} \in \Xi(\cdot)} \sum_d N_d^D \left[\sum_t \left(\sum_g \bar{C}_{dt}^G p_{gdt}^G + \sum_l C_l^{LS} p_{ldt}^{LS} - \sum_c C_{dt}^E p_{cdt}^C \right) \right] \quad (3.14a)$$

Subproblem (3.14) is a bilevel model whose lower-level problem is continuous and linear on its decisions variables. Thus, the lower-level problem can be replaced by its dual constraints. Moreover, using the strong duality equality, we can replace the objective function of the subproblem by the dual lower-level objective function. As a result, the subproblem is finally formulated as the following single-level problem:

$$\max_{\Phi^{L2}, \Phi_{dt}^{L3}}$$

$$\begin{aligned} & \sum_d \left\{ \sum_t \left[\sum_n \lambda_{ndt}^P \sum_{l \in \Psi_n^L} K_{ldt}^D \bar{p}_l^D + \sum_n \lambda_{ndt}^Q \sum_{l \in \Psi_n^L} \alpha_{ldt} K_{ldt}^D \bar{p}_l^D + \sum_n (V_n \underline{\mu}_{ndt}^N - \bar{V}_n \bar{\mu}_{ndt}^N) + \right. \right. \\ & \quad + \sum_g (\underline{P}_g^G \underline{\mu}_{gdt}^P - \bar{P}_g^G \bar{\mu}_{gdt}^P + \underline{Q}_g^G \underline{\mu}_{gdt}^Q - \bar{Q}_g^G \bar{\mu}_{gdt}^Q) - \sum_l K_{ldt}^D \bar{p}_l^D \bar{\mu}_{ldt}^D - \sum_r K_{rdt}^R \bar{P}_r^R x_r^{R(v)} \bar{\mu}_{rdt}^R \\ & \quad - \sum_s \bar{E}_s^S x_s^{S(v)} \bar{\mu}_{sdt}^E + \sum_s \underline{E}_s^S x_s^{S(v)} \underline{\mu}_{sdt}^E - \sum_s \bar{P}_s^{SC} x_s^{S(v)} \bar{\mu}_{sdt}^{SC} \\ & \quad - \sum_s \bar{P}_s^{SD} x_s^{S(v)} \bar{\mu}_{sdt}^{SD} - \sum_c \bar{P}_c^C x_c^{C(v)} \bar{\mu}_{cdt}^C - \sum_c \bar{n}_{cd}^{EV} P_{cd}^{EV} K_{cdt}^{EV} \bar{\mu}_{cdt}^{EV} \left. \right] \\ & \quad \left. - \sum_c \bar{n}_{cd}^{EV} P_{cd}^{EVmax} \varphi_{cd}^{EV} \right\} \end{aligned} \quad (3.15a)$$

subject to

$$\text{Constraints (3.6)} \quad (3.15b)$$

$$N_d^D \bar{c}_{dt}^G = \lambda_{n(g)dt}^P - \bar{\mu}_{gdt}^P + \underline{\mu}_{gdt}^P, \quad \forall g, \forall d, \forall t, \quad (3.15c)$$

$$0 = \lambda_{n(g)dt}^Q - \bar{\mu}_{gdt}^Q + \underline{\mu}_{gdt}^Q, \quad \forall g, \forall d, \forall t, \quad (3.15d)$$

$$N_d^D C_l^{LS} = \lambda_{n(l)dt}^P + \alpha \lambda_{n(l)dt}^Q - \bar{\mu}_{ldt}^D + \underline{\mu}_{ldt}^D, \quad \forall l, \forall d, \forall t, \quad (3.15e)$$

$$-N_d^D C_{dt}^E = -\lambda_{n(c)dt}^P + \underline{\mu}_{cdt}^C - \bar{\mu}_{cdt}^C + \underline{\mu}_{cdt}^{EV} - \bar{\mu}_{cdt}^{EV} - \varphi_{cd}^{EV}, \quad \forall c, \forall d, \forall t, \quad (3.15f)$$

$$-\lambda_{s(\ell)dt}^P + \lambda_{r(\ell)dt}^P + \frac{\kappa_l}{V_0} \varphi_{ldt}^L = 0, \quad \forall \ell, \forall d, \forall t, \quad (3.15g)$$

$$-\lambda_{s(\ell)dt}^Q + \lambda_{r(\ell)dt}^Q + \frac{\chi_l}{V_0} \varphi_{ldt}^L = 0, \quad \forall \ell, \forall d, \forall t, \quad (3.15h)$$

$$\lambda_{n(r)dt}^P + \lambda_{n(r)dt}^Q + \underline{\mu}_{rdt}^R - \bar{\mu}_{rdt}^R = 0, \quad \forall r, \forall d, \forall t, \quad (3.15i)$$

$$\lambda_{n(s)dt}^P \eta_s^{SD} + \varphi_{sdt}^E \Delta t + \underline{\mu}_{sdt}^{SD} - \bar{\mu}_{sdt}^{SD} = 0, \quad \forall s, \forall d, \forall t, \quad (3.15j)$$

$$-\lambda_{n(s)dt}^P - \varphi_{sdt}^E \eta^{SC} \Delta t + \underline{\mu}_{sdt}^{SC} - \bar{\mu}_{sdt}^{SC} = 0, \quad \forall s, \forall d, \forall t, \quad (3.15k)$$

$$\varphi_{sdt}^E - \varphi_{sd,t-1}^E + \underline{\mu}_{sdt}^E - \bar{\mu}_{sdt}^E = 0, \quad \forall s, \forall d, \forall t \quad (3.15l)$$

$$\underline{\mu}_{ndt}^N - \bar{\mu}_{ndt}^N - \sum_{\ell|s(\ell)=n} \varphi_{ldt}^L + \sum_{\ell|r(\ell)=n} \varphi_{ldt}^L = 0, \quad \forall d, \forall t, \forall n \setminus n : \text{ref.}, \quad (3.15m)$$

$$\underline{\mu}_{ndt}^N - \bar{\mu}_{ndt}^N - \sum_{\ell|s(\ell)=n} \varphi_{ldt}^L + \sum_{\ell|r(\ell)=n} \varphi_{ldt}^L + \xi_{ndt} = 0, \quad \forall d, \forall t, n : \text{ref.}, \quad (3.15n)$$

$$\underline{\mu}_{ndt}^N, \bar{\mu}_{ndt}^N \geq 0, \quad \forall n, \forall d, \forall t, \quad (3.15o)$$

$$\underline{\mu}_{gdt}^P, \bar{\mu}_{gdt}^P \geq 0, \quad \forall g, \forall d, \forall t, \quad (3.15p)$$

$$\underline{\mu}_{gdt}^Q, \bar{\mu}_{gdt}^Q \geq 0, \quad \forall g, \forall d, \forall t, \quad (3.15q)$$

$$\underline{\mu}_{ldt}^D, \bar{\mu}_{ldt}^D \geq 0, \quad \forall l, \forall d, \forall t, \quad (3.15r)$$

$$\underline{\mu}_{rdt}^R, \bar{\mu}_{rdt}^R \geq 0, \quad \forall r, \forall d, \forall t, \quad (3.15s)$$

$$\underline{\mu}_{sdt}^E, \bar{\mu}_{sdt}^E \geq 0, \quad \forall s, \forall d, \forall t, \quad (3.15t)$$

$$\underline{\mu}_{sdt}^{SC}, \bar{\mu}_{sdt}^{SC} \geq 0, \quad \forall s, \forall d, \forall t, \quad (3.15u)$$

$$\underline{\mu}_{sdt}^{SD}, \bar{\mu}_{sdt}^{SD} \geq 0, \quad \forall s, \forall d, \forall t, \quad (3.15v)$$

$$\underline{\mu}_{cdt}^C, \bar{\mu}_{cdt}^C \geq 0, \quad \forall c, \forall d, \forall t, \quad (3.15w)$$

$$\underline{\mu}_{cdt}^{EV}, \bar{\mu}_{cdt}^{EV} \geq 0, \quad \forall c, \forall d, \forall t, \quad (3.15x)$$

$$\varphi_{cd}^{EV} \geq 0, \quad \forall c, \forall d. \quad (3.15y)$$

3.3.3 Algorithm

Given the master problem (3.11) and subproblem (3.13), the column-and-constraint generation algorithm works as follows:

1. Initialize the iteration counter ($v \leftarrow 0$), select the convergence tolerance (ϵ), and set the lower

bound (LB) and upper bound (UB) to $-\infty$ and $+\infty$, respectively.

2. Solve master problem (3.11).
3. Update the lower bound using equation (3.16) below:

$$\text{LB} = z^{\text{M}*}, \quad (3.16)$$

where $z^{\text{M}*}$ is the optimal value of the objective function (3.11a).

4. Set $\mathbf{x}^{(v)} = \mathbf{x}^*$, where $\mathbf{x}^*$ denotes the optimal values of variables in vector $\mathbf{x}$ yielded by the solution of the master problem (3.11) in Step 2.
5. Solve subproblem (3.13).
6. Update the upper bound using equation (3.17) below:

$$\text{UB} = \min\{\text{UB}, \mathbf{C}^T \mathbf{x} + z^{\text{S}*}\}, \quad (3.17)$$

where $z^{\text{S}*}$ is the optimal value of the objective function (3.13).

7. If $\text{UB} - \text{LB} < \epsilon$, the algorithm stops. The optimal expansion decisions are $\mathbf{x}^*$. Otherwise, go to step 8.
8. Update the iteration counter $v \leftarrow v + 1$.
9. Set $\mathbf{v}^{(v)} = \mathbf{v}^*$, where $\mathbf{v}^*$ denotes the optimal values of variables in vector $\mathbf{y}$ yielded by the solution of the subproblem (3.13) in Step 5.
10. Go to step 2.

3.4 Case Study

The proposed ARO model is analyzed using the modified 69-bus distribution test system (see Baran and Wu (1989)) depicted in Figure 3.2. Node 1 is the reference node and is also connected with the main grid. Its voltage magnitude is set equal to 1 p.u., while the minimum and maximum voltage magnitudes applied to the other nodes are 0.9 p.u. and 1.1 p.u., respectively. The network technical data are reported in Table 3.A.1 in 3.A. In this distribution system, residential demand load is located

in 48 nodes that are indicated by arrows and can be served using either electricity produced by the PV, or provided by storage units, or withdrawn from the main grid with which node 1 is connected. Candidate nodes for investment in charging stations, PV units, and storage units are respectively indicated with “C”, “R”, and “S” in Figure 3.2. Charging stations can be built in nodes 34 and 61, PV power plants in nodes 24 and 61, and storage units in nodes 12 and 46, respectively. Up to 10000 candidates storage units, 10000 candidate PV plants, and 10000 candidate charging stations are considered as expansion planning options. The number of candidate units is large enough so that the expansion options are not limited by this number, and the optimal expansion plan is determined by the proposed model and the available investment budget. Table 3.1 reports technical data for these units. Each candidate charging station has a capacity of 3.3 kW and an installation cost of \$3000 (Smith and Castellano (2015)). Similarly, each commercial rooftop PV unit has a capacity of 10 kW and an investment cost of \$21400 (Fu et al. (2017)). Each lithium-ion battery unit is assumed to have a power capacity equal to 10 kW, i.e., each unit can charge/discharge at maximum 10 kWh within an hour, while the energy storage capacity is set to 100 kWh. Charging η_s^{SC} and discharging η_s^{SD} efficiencies, as used in 3.6, are set equal to 0.9. Following Pandžić et al. (2015), the investment cost for each 10-kW unit is set equal to \$7000. Recall that investment costs reported in Table 3.1 are annualized in the objective function (3.8a) to make them comparable with operations costs. In the analysis below, a 10% rate for the annualization of these investment costs is considered.

Candidate unit	Capacity [kW]	Investment Cost [\$]
Storage unit	10	7000
PV plant	10	21400
Charging station	3.3	3000

Table 3.1: Data for candidate units.

Technical features for EVs are taken from Sun et al. (2015): energy rating is set to 0.33 kWh/mi, battery capacity is set to 24 kWh and maximum charging rate is set equal to 3.3 kW. All EVs can charge immediately after the arrival, i.e., no smart charging strategy is considered. The possible departure and arrival locations of EVs are “home” (identified by node 34) and “work” (corresponding to node 61), where these EVs can be charged by connecting to the dedicated stations.

As explained in Section 3.2.2, a set of representative days is considered to characterize short-term variability in the demand, the production of PV units, the electricity price exchanged with the main grid, and the EV driving requirements.

Data from the 2009 National Household Travel Survey (NHTS) for the Texas State are considered

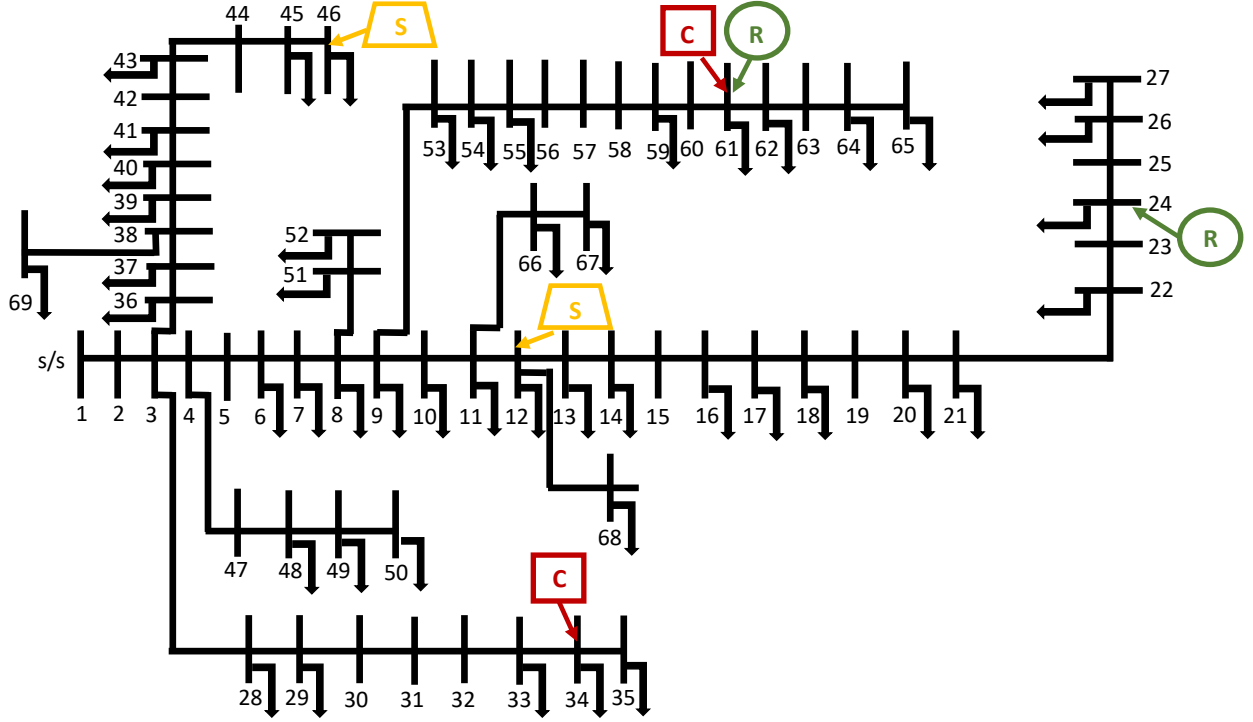


Figure 3.2: 69-bus system

to characterize EV travel behavior, i.e., driver identity, duration, destination, departure and arrival time, length of the trip, and day of the week¹ This dataset refers to fuel vehicles but the typical behavior of EV drivers is assumed to be approximately similar..

Since the vehicle paths change according to the days in a week, these data are classified into two categories: weekend and working days. For each of these two categories, data for “departure”, “arrival”, and “time of the trip” of each single vehicle are combined, obtaining data for parking time in both “work” and “home” locations. EVs are assumed to be always connected and available for charging when parked, i.e., EVs can be charged until the battery is full or they leave for another location. The percentage of battery discharging during the trip is estimated by considering the “miles run” data. In this way, the EV electricity demand for each EV during a 24 hour horizon is obtained. Data are grouped in order to obtain four different hourly EV demand profiles, i.e., “work” and “home” during week days and weekends. Then, each profile is divided by the respective EV peak demand to obtain factors $K_{c dt}^{EV}$. Moreover, parameter $P_{c dt}^{EVmax}$ is also obtained, which limits the daily EV demand considering that the battery has to stop charging when full.

¹Data available at <https://nhts.ornl.gov/>

To maintain correlation among loads, electricity prices, and solar availabilities, historical data used to obtain representative days are all from the same year (2015).

Historical load data are obtained from the Electric Reliability Council of Texas (ERCOT)². More specifically, the historical distribution of loads of the North Central Zone of Texas is considered, which is the area with the highest demand. Loads are then scaled to obtain reference loads for the system. The considered average peak load values $\tilde{P}_l^D$ for the system are reported in Table 3.A.2 in 4.3.

To determine the price of the power exchanged with the main grid, historical day-ahead electricity market prices for the ERCOT North Central load zone are considered.

Solar availabilities are obtained from version 2017.1.17 of the System Advisor Model (SAM)³, based on weather data from the National Solar Radiation Database (NSRDB)⁴. Sengupta et al. (2014,?) validate the NSRDB data and provide details. To maintain the correlation between solar and load data, PV power production data also refer to the North Central Zone of Texas and are obtained as the average production of 12 different locations within the considered zone. This allows obtaining a power production profile that is not influenced by extremely favorable/disadvantageous solar conditions pertaining specific plants.

Once the full data set of loads, electricity price, solar-availability data, and EV driving profiles are defined, the K-means clustering technique explained in Section 3.2.2 is implemented to reduce this full year into a set of representative operating days. For the analysis below, the K-means clustering algorithm is applied to provide 10 clusters, each of which gives one representative day for the distribution planning model, i.e., 10 representative days and a total of 240 operating conditions ($24 \times 10 = 240$) are considered. Further increasing the number of representative days does not have a significant impact on the expansion decisions. The 10 representative days are further subdivided into two categories: working days and weekends. Each representative day has a weight assigned during the clustering procedure. The total weight of the subset of representative days pertaining working days is equal to the number of working days in a year, i.e., 261. Instead, the total weight of the representative days related to weekends is 104, namely the global number of Saturdays and Sundays in a year. This differentiation is introduced to deal with the different time of travel and arrival/departure profiles of EVs during working days and weekends.

To model the long-term uncertainty, the price of the electricity exchanged with the main grid, the

²Data available at www.ercot.com

³Data available at <https://sam.nrel.gov/>

⁴Data available at https://rredc.nrel.gov/solar/old_data/nsrdb/

peak demand and the number of EVs are allowed to deviate up to 25%, 50%, and 25% in the long term with respect to their forecast values, respectively.

The demand uncertainty budget Λ^D can range from 0 to 48, i.e., the number of nodes where demand is located. For the same principle and since there is a single connection to the main grid, the grid uncertainty budget Λ^G can range from 0 to 1, while the EV uncertainty budget Λ^{EV} can range from 0 to 2 because charging stations can be installed at most in two nodes.

Parameter C_{dt}^E , which identifies the revenues that the central planner gets from selling electricity to EVs, is assumed to be twice (unless otherwise specified) the forecast price of the power exchanged with the main grid $\tilde{C}_{dt}^G$. The latter ranges from a minimum of \$13/kWh to a maximum of \$203/kWh according to the operating condition and hour of the day. Note that, in most of the analyzed cases illustrated in the following section, the central planner withdraws electricity from the grid. Finally, load shedding cost is set equal to \$1000/kWh.

3.5 Results

This section provides the results of the case study. Taking as reference the network represented in Figure 3.2 and the input data described in Section 3.4, firstly a set of sensitivity cases is carried out to evaluate how investment decisions change depending on the considered long-term uncertainty (see Section 3.5.1), the revenues of the central planner accruing from selling electricity to EVs (see Section 3.5.2), the degree of independence of the distribution system from the main grid (see Section 3.5.3), and the investment budget availability (see Section 3.5.4). Then, an ex-post estimation of the amount of CO₂ emissions saved thanks to the penetration of renewable energy and EVs is provided (see Section 3.5.5). In addition, Section 3.5.6 investigates the impacts on investment decisions when a larger number of candidate nodes is available for the construction of charging stations, PV power plants, and storage units with respect to that used in the analyses conducted in Sections 3.5.1-3.5.5. Finally, Section 3.5.7 compares the results of the proposed ARO model with those obtained by implementing a corresponding stochastic programming formulation of the model where long-term uncertainty is modeled using scenarios rather than uncertainty sets based on confidence bounds. As explained in this section, this evaluation is done by evaluating the computational performance of the two optimization techniques applied in the presence of uncertainty.

Note that the analysis of each impact listed above entails some modifications of the general ARO model presented in Section 3.2 as explained in the following.

In all analyses, results are obtained using CPLEX (IBM ILOG CPLEX, 2020) under GAMS (Rosenthal, 2012) on a MacBook Pro computer with processor Intel Core i7 at 2.7 GHz and 16 GB of RAM.

3.5.1 Impact of the long-term uncertainty

This section investigates how the investment choices change when setting different values of the uncertainty budgets that regulate the worst possible realizations of long-term uncertain parameters. This evaluation is conducted considering the following assumptions:

1. No limits on the active and reactive power exchanged with the main grid are imposed, namely constraints (3.10f) and (3.10g) are relaxed (see 3.2.6). This case is referred to as “N.L.M.G.”, namely, “no limits, main grid”.
2. Limits on the active and reactive power exchanged with the main grid are imposed, i.e., constraints (3.10f) and (3.10g) hold. More precisely, it is assumed that parameters $\underline{P}^G = \underline{Q}^G = -2800$ (kW, kvar) and $\overline{P}^G = \overline{Q}^G = 2800$ (kW, kvar). This case is referred to as “L.M.G.”, namely, “limits, main grid”. This case is analyzed to detect the level of dependence of the distribution system on the main grid.
3. No limits are enforced on investment budget availability, i.e., constraint (3.18e) is relaxed.

In the “L.M.G.” case, parameters $\overline{P}^G$ and $\overline{Q}^G$ have been defined as follows. Firstly, the ARO model is run considering the “N.L.M.G.” assumptions; obtaining that the average amount of power purchased from the main grid is around 3200 (kW, kvar). No power is sold back to the main grid. Since the “L.M.G.” case aims at representing a distribution system that is already under stress due to a relatively significant restriction imposed on the access to the main grid, it is assumed that $\overline{P}^G = \overline{Q}^G = 2800$ (kW, kvar). For symmetry, it is also considered that $\underline{P}^G = \underline{Q}^G = -2800$ (kW, kvar).

The analysis is started by setting Λ^G , Λ^L , and Λ^{EV} equal to zero, meaning that long-term uncertainty in the peak demand, the value of the electricity exchanged with the main grid, and the number of EVs is not considered and that the model reduces to the deterministic form described in 3.2.3. Note that this directly descends from the construction of uncertainty set Ω defined by constraints (3.6a)-(3.6f). Table 3.2 reports the results for the deterministic case.

In the “N.L.M.G.” case, PV and storage units are not built since the demand of consumers and EVs can be directly satisfied by the electricity withdrawn from the main grid. The central planner only

Deterministic	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
$\Lambda^G = \Lambda^L = \Lambda^{EV} = 0$	91	81	0	90	0	790	0	0.332	427,398	484,761

Table 3.2: Investment decisions for the deterministic case

invests in charging stations because they represent a source of revenues. The situation changes under the “L.M.G.” case. The limited access to the main grid forces the central planner to build 90 kW of solar power plants to satisfy the demand and 790 kW of storage capacity to accommodate the short-term variability. Investments in charging stations remain even though the number of charging stations built is reduced with respect to the “N.L.M.G.” case. This is due to the fact that the limited access to the main grid gives priority to the consumers, i.e., the electricity drained from the main grid is firstly used to satisfy residential demand, as there is a penalty in case of unserved demand, and then the excess is sold to the EV owners. However, this does not suffice since a load shedding of 0.332 kW is registered. As expected, the limited access to the main grid increases the total costs of the distribution system, implying a reduction of its efficiency. Note that in both cases, electricity is only withdrawn from the main grid and it is not sold back.

Λ^G	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
0	72	80	0	1710	0	750	0	104.139	533,962	947,696
1	80	24	0	1750	0	800	0	96.43	678,415	1,064,651

Table 3.3: Impact of the long-term uncertainty of the value of the electricity from the main grid on investment decisions

As soon as the values the uncertainty budget Λ^G , Λ^L , or Λ^{EV} differ from zero, long-term uncertainties are considered. Table 3.3 illustrates the changes in the optimal expansion decisions for different values of the uncertainty budget Λ^G . For this analysis, uncertainty budgets are fixed to $\Lambda^L = 24$ and $\Lambda^{EV} = 1$, meaning that the peak load can increase in up to 24 nodes, representing the worst possible outcomes for the system, and the number of EVs can decrease in one location. When $\Lambda^G = 0$, it is assumed that the value of the electricity exchanged with the main grid does not deviate from its average in the long-run, i.e., it is only considered the effect of the uncertainty in the peak demand and EVs. The obtained results differ from those reported in Table 3.2.

When the distribution system has free access to the main grid (see “N.L.M.G.” columns in Table 3.3), there are only investments in charging stations. Demands of consumers and EVs are satisfied

by withdrawing electricity from the main grid and there is no need to build PV and storage units. This happens both when $\Lambda^G = 0$ and $\Lambda^G = 1$.

On the contrary, the situation changes when the access to the main grid is constrained (see “L.M.G.” columns in Table 3.3). In additions to charging stations, PV units are installed and the storage units are built to compensate the variability of renewable units. However, this does not suffice to cover the demand in all representative days and load shedding is registered. This situation is exacerbated when $\Lambda^G = 1$, i.e., when the price of the electricity exchanged with the main grid can increase in the long run. This leads to two phenomena: first, a significant drop of investments in charging stations due to the reduction of the revenues generated from selling electricity; second, an increase of the amount of PV units installed to make the distribution system more independent of the main grid. This has a positive effect on the load shedding that reduces with respect to the level observed in the $\Lambda^G = 0$ of the “L.M.G.” case. However, this investment strategy leads to comparatively higher total costs. In general, total costs increase compared to the ones obtained in the corresponding deterministic cases. This particularly holds for the cases with limited access to the main grid.

The optimal expansion decisions for different values of the uncertainty budget Λ^L are reported in Table 3.4. In this case, it is assumed that $\Lambda^G = 1$ and $\Lambda^{EV} = 1$. Results are similar to those observed for the analysis of the uncertainty budget Λ^G and reported in Table 3.3. In all “N.L.M.G.” cases, there are only investments in charging stations and electricity demand is completely satisfied by the power taken from the main grid.

PV power plants and storage units are built when the “L.M.G.” assumption applies. More precisely, investments in PV and storage units increase in parallel to the increment of the uncertainty in demand. This is accompanied by a progressive reduction of the investments in charging stations. This investment strategy is applied both to decrease the electricity demand of the system, as the access to the main grid is limited, and to mitigate the load shedding that remains an issue in all “L.M.G.” cases. The higher the level of uncertainty in the peak demand is, the higher the system costs are. This happens both in the “N.L.M.G.” and “L.M.G.” cases. However, the system costs in the “L.M.G.” cases are much higher than in the “N.L.M.G.” and deterministic cases because of the larger investment requirements and the load shedding.

Table 3.5 reports the results for different values of the uncertainty budget Λ^{EV} . In this case, it is considered that $\Lambda^G = 1$ and $\Lambda^L = 2$. As it can be noticed, the variation of Λ^{EV} has a limited impact on the investment strategies and on the total costs, which however remain higher than in the deterministic case.

In all these cases, withdrawal and purchase of electricity represent the net position that the central

Λ^L	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
0	62	56	0	90	0	790	0	0.332	553,396	600,569
12	69	35	0	1380	0	660	0	67.51	661,477	944,994
24	80	24	0	1750	0	800	0	96.43	678,415	1,064,651
36	62	22	0	1870	0	820	0	125.03	687,680	1,122,043
48	62	22	0	1880	0	800	0	136.35	693,492	1,135,679

Table 3.4: Impact of the long-term uncertainty of peak demand on investment decisions

planner has with respect to the main grid.

Λ^{EV}	# Charging stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
0	80	24	0	1750	0	800	0	96.43	678,415	1,064,651
1	80	24	0	1750	0	800	0	96.43	678,415	1,064,651
2	76	24	0	1750	0	800	0	96.43	678,733	1,064,651

Table 3.5: Impact of the long-term uncertainty of the number of EVs on investment decisions

Finally, Table 3.6 reports the execution time of the simulations related to the analyses conducted in this section. Note that the CPU times of the simulations conducted in the following sections are of the same order of magnitude of those reported in Table 3.6. These CPU times are acceptable for the practical requirements of a planning model.

3.5.2 Impact of the revenues associated with EVs

This section evaluates how the investment choices of the central planner change depending on the revenues that it gets from supplying electricity to EVs through charging stations. The following assumptions are taken into account:

1. Both the “N.L.M.G.” and “L.M.G.” cases are considered.
2. Uncertainty budgets are fixed to $\Lambda^G = 1$, $\Lambda^L = 24$, and $\Lambda^{EV} = 1$.
3. No limits are enforced on investment budget availability.

		N.L.M.G.	L.M.G.
Deterministic	$\Lambda^G = 0, \Lambda^L = 0, \Lambda^{EV} = 0$	6 s	15 s
Long-term uncertainty Λ^G	$\Lambda^G = 0, \Lambda^L = 24, \Lambda^{EV} = 1$	32 min	37 min
	$\Lambda^G = 1, \Lambda^L = 24, \Lambda^{EV} = 1$	38 min	42 min
Long-term uncertainty Λ^L	$\Lambda^G = 1, \Lambda^L = 0, \Lambda^{EV} = 1$	40 s	1 min
	$\Lambda^G = 1, \Lambda^L = 12, \Lambda^{EV} = 1$	34 min	39 min
	$\Lambda^G = 1, \Lambda^L = 24, \Lambda^{EV} = 1$	38 min	42 min
	$\Lambda^G = 1, \Lambda^L = 36, \Lambda^{EV} = 1$	47 min	55 min
	$\Lambda^G = 1, \Lambda^L = 48, \Lambda^{EV} = 1$	61 min	78 min
Long-term uncertainty Λ^{EV}	$\Lambda^G = 1, \Lambda^L = 24, \Lambda^{EV} = 0$	36 min	40 min
	$\Lambda^G = 1, \Lambda^L = 24, \Lambda^{EV} = 1$	38 min	42 min
	$\Lambda^G = 1, \Lambda^L = 24, \Lambda^{EV} = 2$	41 min	45 min

Table 3.6: Execution time of the long-term uncertainty analysis with the ARO approach

In general, charging stations are built only if they represent a profit for the central planner, i.e., when the difference between the revenues obtained from selling electricity to EV owners through charging stations and the electricity provision costs are high enough to cover part or all investment costs in charging stations. It is important to recall that the electricity provided to the charging stations derives from either the main grid, or the PV units, or the storage units. Since the electricity from PV units is assumed to be freely generated, it becomes crucial the margin between the price applied to the electricity sold to the EV owners C_{dt}^E and the forecast price of the power exchanged with the main grid $\tilde{C}_{dt}^G$. For this reason, a sensitivity analysis is conducted on the values assumed for parameter C_{dt}^E compared to $\tilde{C}_{dt}^G$. The range of values is reported in Table 3.7.

Values of C_{dt}^E	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
$C_{dt}^E = 1.5\tilde{C}_{dt}^G$	0	0	0	1730	0	790	0	96.17	687,839	1,061,237
$C_{dt}^E = 1.8\tilde{C}_{dt}^G$	26	0	0	1740	0	790	0	98.35	686,221	1,063,903
$C_{dt}^E = 2\tilde{C}_{dt}^G$	76	24	0	1750	0	800	0	96.43	678,733	1,064,651
$C_{dt}^E = 3\tilde{C}_{dt}^G$	83	88	0	1750	0	800	0	96.66	640,880	1,032,438
$C_{dt}^E = 5\tilde{C}_{dt}^G$	138	117	0	1790	0	810	0	94.41	539,496	954,554
$C_{dt}^E = 10\tilde{C}_{dt}^G$	217	186	0	1960	0	790	0	80.07	307,726	743,473

Table 3.7: Impact of the revenues associated with EVs on investment decisions

Considering the “N.L.M.G” cases, it is observed that when the value of C_{dt}^E is 50% higher than that of $\tilde{C}_{dt}^G$ there is no installation of charging stations, which start to be built when this gap is equal to 80%, i.e., $C_{dt}^E = 1.8\tilde{C}_{dt}^G$. From this point on, there is a progressive increasing trend in investments in charging units as far as C_{dt}^E doubles the value of $\tilde{C}_{dt}^G$. Under the assumption of free access to the main grid, there are no investments in PV and storage units independently of the

value of C_{dt}^E . It is remarkable the progressive reduction of the total costs of the system that derives from the increasing revenues that the central planner obtains when selling electricity to the EVs through the charging stations. However, these revenues do not suffice to cover all the investments and operating costs faced by the central planner. More precisely, the electricity provided by the PV power plants built and the storage units, when discharging, does not cover all the demand of consumers and EVs and, therefore, the distribution system has to withdraw electricity from the main grid. Note that the amount of electricity required by local consumers is proportionally higher than the amount needed by the EVs.

The situation partially changes when the “L.M.G” cases are analyzed. When the relations $C_{dt}^E = 1.5\tilde{C}_{dt}^G$ and $C_{dt}^E = 1.8\tilde{C}_{dt}^G$ hold, it is not convenient investing in charging stations. However, due to the limited access to the main grid, PV units are needed to supply demand. These power plants are built together with storage units that are used to provide more flexibility to the system. Investments in charging stations appear when C_{dt}^E doubles the value of $\tilde{C}_{dt}^G$ and then start increasing for higher ratios between the two parameters. As also observed for the “N.L.M.G” cases, the raise of the revenues from EVs implies a reduction of the total costs. There is load shedding in all “L.M.G” scenarios; the lowest amount of unserved demand is registered in when $C_{dt}^E = 10\tilde{C}_{dt}^G$ applies because of the comparatively higher availability of PV power plants.

3.5.3 Impact of the degree of dependency of distribution system from the main grid

This section provides a deeper analysis of the impact that the dependency of the distribution system from the main grid has on the expansion decisions. The illustrated results are based on these assumptions:

1. The “L.M.G.” case is considered as reference.
2. Uncertainty budgets are fixed to $\Lambda^G = 1$, $\Lambda^L = 24$, and $\Lambda^{EV} = 1$.
3. No limits are enforced on investment budget availability.

In particular, since in all the previous simulations the system is only importing power from the main grid, here it is analyzed how expansion decisions change when the maximum amount of electricity drainable from the main grid progressively reduces. For this reason, starting from the reference average value of 3200 (kW, kvar), it is assumed that parameters $\bar{P}^G$ and $\bar{Q}^G$ can take the

values of 3000 (kW, kvar), 2800 (kW, kvar), 2500 (kW, kvar), and 2000 (kW, kvar), corresponding to a reduction of 6%, 13%, 22%, and 38% of the reference amount withdrawn, respectively, while the lower bounds $\underline{P}^G$ and $\underline{Q}^G$, i.e., the minimum hourly amount of electricity that can sold to the grid, are fixed to -3000 (kW, kvar), -2800 (kW, kvar), -2500 (kW, kvar), and -2000 (kW, kvar), respectively.

$\overline{P}^G, \overline{Q}^G, \underline{P}^G, \underline{Q}^G$	# Charging Stations	Solar Capacity (kW)	Storage Capacity (kW)	Load Shedding (kW)	Total Cost (\$)
No limits	80	0	0	0	678,733
3000 (kW, kvar)	45	1270	690	44.89	934,997
2800 (kW, kvar)	24	1750	800	96.43	1,064,651
2500 (kW, kvar)	22	2350	1000	361.92	1,424,673
2000 (kW, kvar)	0	2400	1000	1,833.89	2,856,514

Table 3.8: Impact of the limits of the power purchased from main grid on investment decisions

The results of this analysis are reported in Table 3.8, where row “No limits” reports the expansion decisions taken by the central planner when it can freely withdraw electricity from the main grid. When the supply from the main grid is restricted, the central planner prefers to invest more in PV power plants and storage units, while reducing the number of charging stations built. This is a direct consequence of the forced limited access to the main grid: additional PV plants are needed to cover the electricity demand and storage units are built to mitigate solar production variability. Moreover, as already observed, limiting the number of charging stations helps reducing the electricity demand of the system. However, this does not suffice to satisfy the whole demand requirement since there is a progressive increase of load shedding that reaches the value of 1833.89 kW when the access to the main grid is limited to 2000 (kW, kvar). Note that in this case, the central planner invests in PV and storage units, but not in charging stations for the reasons explained above. With this configuration, the system is more autonomous but its management becomes more expensive leading to increasing total costs as soon as the access to the main grid becomes more limited. Finally, in all these cases, withdrawal and purchase of electricity represent the net position that the central planner has with respect to the main grid.

3.5.4 Impact of the investment budget availability

This section evaluates how the budget availability can influence investment decisions. This analysis is based on the following assumptions:

1. The “L.M.G.” case is considered as the reference.
2. Uncertainty budgets are fixed to $\Lambda^G = 1$, $\Lambda^L = 24$, and $\Lambda^{EV} = 1$.

The results of this analysis are reported in Table 3.9. Firstly, the ARO model is run considering the assumptions listed above and assuming that no budget limits apply. This model configuration leads to the “No limits” results reported in Table 3.9 and shows that the obtained investment mix requires an annualized investment budget of \$437,700. Taking as reference this value, we construct a set of sensitivity cases where it is assumed that the annualized investment budget is reduced by 25%, 50%, and 75% leading to a budget availability of \$328,275, \$218,850, and \$109,425, respectively.

Budget	# Charging Stations	Solar Capacity (kW)	Storage Capacity (kW)	Load Shedding (kW)	Investment Cost (\$)	Operating Cost (\$)	Total Cost (\$)
No limits	24	1750	800	96.43	437,700	626,951	1,064,651
\$328,275	0	1320	650	190.88	327,980	756,715	1,084,695
\$218,850	0	810	650	335.36	218,840	929,933	1,148,773
\$109,425	0	310	610	663.29	109,040	1,278,047	1,387,087

Table 3.9: Impact of the budget availability on investment decisions

In the absence of investment budget limits, charging stations, PV power plants, and storage units are built. A restricted budget availability leads to the following consequences: (1) charging stations are no longer constructed; (2) the central planner still invests in PV and storage units; (3) total costs progressively increase due to an increment of the operating costs. With a limited budget availability, the central planner decides to cut investments in PV power plants that is the most expensive technology among the ones considered. The cut in investments in PV leads to an increase of the amount of unserved demand as the distribution system becomes more dependent on the main grid. In addition, the central planner decides to build no charging stations to reduce the electricity demand of the whole distribution system and to contain load shedding. The combination of these actions results in an increase of the operating costs related to the purchase of electricity from the main grid and to load shedding that imply an overall increment of the system cost. This holds for the budget reduction cases considered and the phenomenon is particularly exacerbated when the available budget is very limited, i.e., in the \$109,425-budget case.

3.5.5 An ex-post decarbonization analysis

This section analyzes whether the adoption of EVs can be environmentally advantageous and can contribute to the achievement of the decarbonization targets imposed by the European Commis-

sion. The penetration of EVs leads to a trade-off: on the one side, these are introduced to replace pollutant conventionally fuel cars; on the other side, a progressive integration of EVs would imply an increase of electricity demand. In this case, the main issue is to understand which sources are used for power generation. If EVs are charged using renewable energy only, this would significantly contribute to mitigate CO₂ emissions; otherwise the emissions saved by the replacement of conventionally fuel cars with EVs are partially or, in the worst case, totally compensated by those emitted when producing the electricity that EVs need. Translating this for the considered case study means that the increase of electricity demand due to the penetration of EVs is covered by importing more power from the main grid, or alternatively by installing PV units in the distribution system. On the contrary, a system without EVs can import less power from the main grid, but it emits CO₂ by using conventional fuel vehicles.

This decarbonization analysis is based on the following assumptions:

1. Both the “N.L.M.G.” and “L.M.G.” cases are considered.
2. Uncertainty budgets are fixed to $\Lambda^G = 1$, $\Lambda^L = 24$, and $\Lambda^{EV} = 1$.
3. The revenues associated with EVs are set at $C_{dt}^E = 2\tilde{C}_{dt}^G$.
4. No limits are enforced on investment budget availability.

To detect the possible level of decarbonization achieved with the integration of EVs, two situations are compared: the first corresponds to the formulation discussed in the chapter where EVs circulate and are charged; in the second, it is assumed that the charged EVs are replaced by diesel vehicles. In the following, these two situations are denoted as “EVs” and “NO EVs”, respectively. Then, the emissions generated by these two configurations of the distribution system are computed: in the first one, CO₂ only derives from electricity production; in the second, carbon emissions are a “negative” output of both power generation and car utilization. First, the amount of CO₂ emitted by electricity production is computed. Solar energy generated by PV units is carbon free. However, emissions could be generated by the electricity withdrawn from the main grid that is used to supply the demand of EVs and consumers. As explained in Section 3.4, the input data mainly refer to Texas. For this reason, the fuel mix used to generate electricity in that area of the US is considered. According to the Alternative Fuel Data Center⁵, approximately 45% of the electricity production in Texas is generated with natural gas, 29% from coal, and the rest from CO₂ free sources, including

⁵See https://www.afdc.energy.gov/vehicles/electric_emissions.php

nuclear power. Taking as reference this fuel mix, the electricity taken from main grid is assigned an emission factor of 0.466 ton CO₂/MWh that is computed as a weighted average of the emission factors associated with the fuels used to produce electricity and their corresponding proportion in the Texan power generation.

To compute the carbon emissions of diesel vehicles, it is assumed that:

- Diesel vehicles replace charging EVs, i.e., the number of cars circulating in the system remains the one resulting from the analysis.
- Diesel vehicles replicate exactly the travel path of the EVs.
- 8.5 km/liter and 18.8 km/liter consumption are assumed for diesel vehicles.
- The CO₂ emission rate associated with diesel vehicles is equal to 2.6391 kg CO₂/liter of fuel⁶.

Note that 8.5 km/liter and 18.8 km/liter correspond to 11.76 liters/100km and 5,32 liters/100km, which are the average fuel consumption levels of diesel pickup trucks/vans and cars, respectively⁷.

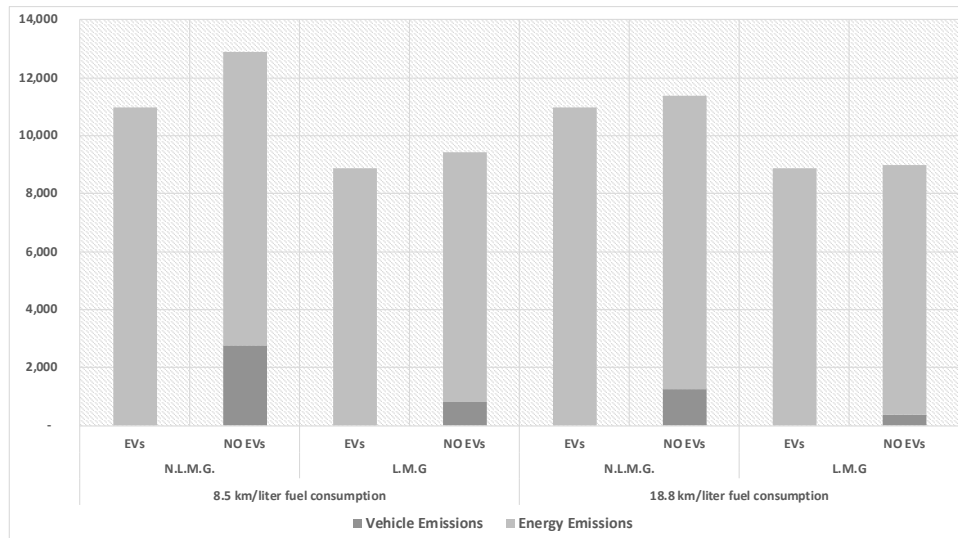


Figure 3.3: Annual ton CO₂ emissions in the considered cases

Figure 3.3 shows the annual amount of CO₂ generated in the considered cases. For each market configurations, the CO₂ emissions generated by the whole system in presence and in absence of

⁶See <https://carbonpositivelife.com/co2-per-litre-diesel/>

⁷For more details, see <https://www.nrcan.gc.ca/2019-fuel-consumption-guide/21002>

EVs are compared. The “EVs” situation considers the emissions generated by the total amount of electricity that supply the system’s demand, including the power consumed by EVs and local consumers. The “NO EVs” situation considers the emissions that derive from the electricity required by local consumers only and CO₂ emitted by the diesel fuelled cars. The electricity needed by local consumers is the same in the two situations; what makes the difference is the power used to supply charging stations that is not accounted for in the “NO EVs” situation.

From Figure 3.3, it is observed that, in all cases, the adoption of EVs allows saving CO₂ emissions compared to the corresponding “NO EVs” situation, notwithstanding the fact that their utilization implies a larger electricity consumption level. Moreover, the total amount of emissions generated in the “L.M.G.” cases is lower than the level registered in the “N.L.M.G.” scenarios. This is in line with the results obtained in the analysis illustrated in Section 5.1: the limited access to the main grid makes the distribution system more independent and induces the central planner to invest in the carbon free PV power plants to cover the demand. As a direct consequence, CO₂ emissions reduce in comparison with the “N.L.M.G.” cases, where the demands of EVs and consumers are covered by electricity withdrawn from the main grid.

When considering a fuel consumptions of 8.5 km/liter, the emission cuts are equal to 15% and 6% in the “N.L.M.G.” and “L.M.G.” cases, respectively. The higher emission reduction observed when the “N.L.M.G.” system configuration applied is due to the fact the number of EVs circulating is larger (see Tables 3.3-3.5 in Section 5.1). This means that the integration of EVs, supported by a parallel penetration of renewable-based power plants and an adequate network development, significantly contribute to climate change mitigation and to the decarbonization targets. These reduction rates are lower when considering a more efficient set of vehicles with a fuel consumption of 18.8 km/liter. In particular, these amounts are 3% and 1% in the “N.L.M.G.” and “L.M.G.” cases, respectively.

3.5.6 Impact of the availability of a larger number of candidate nodes

This section investigates how investment decision change when charging stations, PV power plants, and storage units can be built in more than 2 candidate nodes as considered in the analyses described in the previous sections. In particular, there are 8 candidate nodes for charging stations, namely, nodes 6, 17, 27, 34, 40, 48, 52, and 61; 8 candidate nodes where PV power plants can be built, namely, nodes 12, 17, 21, 24, 48, 50, 61, and 64; and 8 candidate nodes where storage units can be constructed, namely, nodes 12, 21, 24, 28, 34, 46, 48, and 63 (see Figure 3.4). The technical features of all these units are as described in Section 3.4. As for charging stations, candidate nodes

27, 34, 40, and 48 are considered as “home” locations, while candidate nodes 6, 17, 52, and 61 are “work” locations. For the sake of simplicity, in the following, it is denoted as “2CN” and “8CN” the network configurations where there are, respectively, 2 and 8 candidates nodes destined to the investments in charging stations, PV power plants, and storage units.

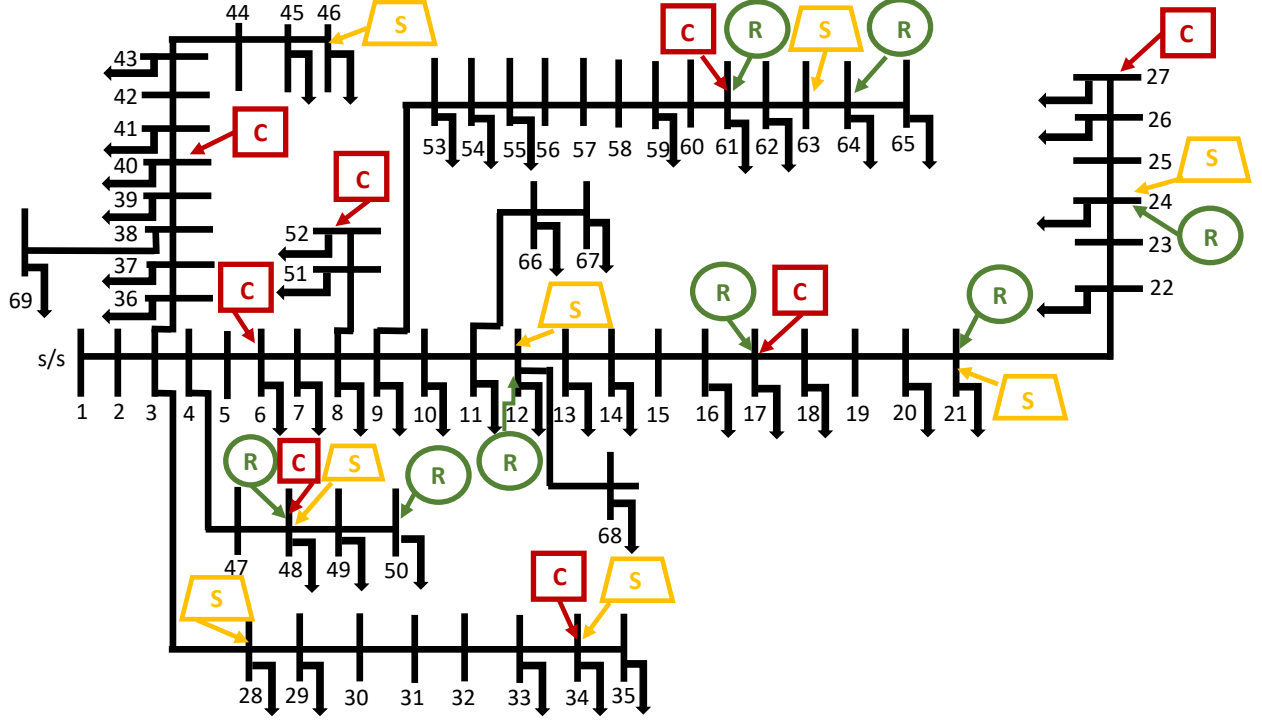


Figure 3.4: 69-bus system under the “8CN” the network configuration assumption

Similarly to what already done in Section 3.5.1, it is analyzed how long-term uncertainty affects investment decisions. This is done by considering the following assumptions:

1. Both the “N.L.M.G.” and “L.M.G.” cases are considered. In order to enable the comparison between the results of the “2CN” and “8CN” network configurations, the same limits on the active and the reactive power exchanged with the main grid are imposed, i.e., $\underline{P}^G = \underline{Q}^G = -2800$ (kW, kvar) and $\overline{P}^G = \overline{Q}^G = 2800$ (kW, kvar).
2. The revenues associated with EVs are set at $C_{dt}^E = 2\tilde{C}_{dt}^G$.
3. No limits are enforced on investment budget availability.

Firstly, the “8CN” network configuration under the assumption $\Lambda^G = \Lambda^L = \Lambda^{EV} = 0$ is considered, which corresponds to the deterministic case. Results are reported in Table 3.10. In the “N.L.M.G.”

Deterministic	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
$\Lambda^G = \Lambda^L = \Lambda^{EV} = 0$	364	155	0	100	0	770	0	0	364,319	468,729

Table 3.10: Results for the deterministic case

case only charging stations are built and the demand of consumers and EVs is covered with the electricity directly taken from the main grid. In the “L.M.G.” case, the limited access to the main grid has several implications on the investment strategies adopted. In particular, the central planner has to construct 100 kW of PV plants to satisfy electricity demand because the restricted access to the main grid forces the distribution system to be more independent. In addition, 770 kW of storage units are built to compensate for the short-term variability of renewable-based power units. There are still investments in charging stations even though in a reduced number compared to the “N.L.M.G.” case. The final effect is that the total system costs in the “L.M.G.” case are higher than under the “N.L.M.G.” assumption. This trend is similar but not identical to that described in Section 3.5.1 for the deterministic case applied to the “2CN” network design. First of all, in the “8CN” network configuration the number of charging stations built is higher than in “2CN” market design thanks to a larger availability of candidate nodes. In contrast, the increased number of candidate nodes does not significantly modify the investment choices in PV power plants and storage units, whose capacity built is almost the same in the two network configurations (compare Tables 3.2 and 3.10). The combination of this investment strategies has a positive impact on the system total costs in the “8CN” network, which are lower compared to those obtained in the corresponding “2CN” case because of the higher revenues accruing from EVs using charging stations (compare Tables 3.2 and 3.10). This holds both under the “N.L.M.G.” and “L.M.G.” assumptions. Second, there is no load shedding in the “L.M.G.” case (compare Tables 3.2 and 3.10).

Λ^L	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
0	284	112	0	100	0	770	0	0	528,117	592,162
12	319	88	0	1450	0	680	0	73.31	631,938	963,555
24	297	88	0	1740	0	800	0	94.44	653,007	1,057,106
36	282	87	0	1850	0	800	0	126.668	662,327	1,116,430
48	290	84	0	1890	0	800	0	139.558	666,261	1,138,776

Table 3.11: Impact of the long-term uncertainty of peak demand on investment decisions

Table 3.11 reports the optimal expansion decisions for different values of the uncertainty budget

Λ^L . Uncertainty budgets Λ^G and Λ^{EV} are both set equal to 1 in order to make the results of the “8CN” and “2CN” network configurations comparable. Considering the “8CN” network design, in all “N.L.M.G.” cases the central planner only invests in charging stations and electricity is withdrawn from the main grid to satisfy the whole system demand. Investments in PV power plants and in storage units are planned in all “L.M.G.” cases and the corresponding amount of capacity built increases proportionally to the increment of the demand uncertainty. In parallel, there is a progressive reduction in the number of charging stations constructed. As explained above, these are direct consequences of the limited exchange with the main grid. Moreover, load shedding appears in the “L.M.G.” cases with the exception of the $\Lambda^G = 1, \Lambda^L = 0, \Lambda^{EV} = 1$ one. Both in the “N.L.M.G.” and “L.M.G.” cases, the number of charging stations built in the “8CN” case is larger than that observed when the “2CN” assumption applies (compare Tables 3.4 and 3.11). However, investments in new PV power plants and storage units are very similar. The combination of these investment decisions again makes the total system costs of all cases in the “8CN” network configuration lower than those registered when the “2CN” assumption is considered.

Λ^{EV}	# Charging Stations		Solar Capacity (kW)		Storage Capacity (kW)		Load Shedding (kW)		Total Cost (\$)	
	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.	N.L.M.G.	L.M.G.
0	320	88	0	1710	0	780	0	94.79	650,305	1,050,109
1	297	88	0	1740	0	800	0	94.44	653,007	1,057,106
2	266	88	0	1740	0	800	0	94.48	654,648	1,057,390
4	248	82	0	1740	0	800	0	94.53	656,469	1,057,419
8	244	68	0	1710	0	780	0	95.34	659,610	1,051,525

Table 3.12: Impact of the long-term uncertainty of the number of EVs on investment decisions

Table 3.12 reports the results obtained for different values of the uncertainty budget Λ^{EV} when considering the “8CN” network configuration. As in Section 3.5.1, it is assumed that $\Lambda^G = 1$ and $\Lambda^L = 24$. Moreover, since the “8CN” case assumes the presence of 8 candidate nodes for the construction of EVs, it is imposed that Λ^{EV} can be either equal to 0, or to 1, or to 2, or to 4, or to 8. The investment choices of the central planner are as described above both in the “N.L.M.G.” and “L.M.G.” cases (compare Tables 3.11 and 3.12). Note that, in the “N.L.M.G.” cases, increasing the uncertainty related to the number of EVs implies a progressive reduction of the number of charging stations constructed and a parallel increase of the total operating costs. This also holds for all the “L.M.G.” cases with the exception of the $\Lambda^{EV} = 8$ instance. In this case, the high level of uncertainty makes the investments in charging stations more risky and therefore the amount of charging stations built

reduces, implying a possible decrease of the power demand from EVs. Therefore, the central planner constructs 1710 kW of PV power units and 780 kW of storage units as in the $\Lambda^G = 1$, $\Lambda^{EV} = 24$, $\Lambda^{EV} = 0$ instance. In this way, the total operating costs result to be lower than in the cases with $\Lambda^{EV} = 1$, $\Lambda^{EV} = 2$, and $\Lambda^{EV} = 4$, but they are higher than in the $\Lambda^{EV} = 0$ instance because the registered load shedding is higher. Finally, the insights provided by the comparison between the results of the “8CN” and “2CN” network configurations are in line with those already analyzed for the long-term uncertainty of peak demand.

3.5.7 Modeling long-term uncertainties using scenarios: a stochastic programming approach

The distribution expansion planning problem is solved in this section using a two-stage stochastic programming problem (see, Birge and Louveaux (1997)) in which long-term uncertainties are modeled using a set of scenarios. In this problem, first-stage decisions are the expansion decisions to be made in the distribution system, namely, the charging stations, storage units, and renewable units to be built, which are here-and-now decisions since they do not depend on the future scenario realizations. On the other hand, second-stage decisions refer to operating decisions and are wait-and-see decisions that depend on the actual scenario realization for the stochastic programming formulation of the distribution expansion planning problem. In particular, the parameters $C_{dt}^G(\omega)$, $P_t^D(\omega)$, and $N_{cd}^{EV}(\omega)$ identify the prices of the power exchanged with the main grid in hour t of the representative day d , the peak load of demand l , and the number of EVs in charging station c in representative day d , respectively. These parameters depend on scenario ω with $\omega \in \Omega^\omega$, where Ω^ω is the set of all scenarios. The parameter $\tau(\omega)$ represents the weight assigned to each scenario ω . First-stage variables are x_s^S , x_r^R , and x_c^C , while the remaining variables are two-stage variables that depend on scenario ω . The meaning of the variables and the parameters are as described in 3.2.1. The model structure is very similar to the corresponding deterministic version presented in 3.2.3. The reader is referred to that section for the explanation of the constraints. Finally, note that the short-term variability is also considered in this two-stage stochastic programming formulation of the model and it is represented using representative days d as in the ARO approach.

$$\min_{\Phi^S} \sum_s C_s^{IS} x_s^S + \sum_r C_r^{IR} x_r^R + \sum_c C_c^{IC} x_c^C$$

$$+ \sum_{\omega} \tau(\omega) \sum_d N_d^D \left[\sum_t \left(\sum_g C_{dt}^G(\omega) p_{gdt}^G(\omega) + \sum_l C_l^{LS} p_{ldt}^{LS}(\omega) - \sum_c C_{dt}^E p_{cdt}^C(\omega) \right) \right] \quad (3.18a)$$

subject to

$$x_s^S \in \mathbb{Z}^+, \quad \forall s, \quad (3.18b)$$

$$x_r^R \in \mathbb{Z}^+, \quad \forall r, \quad (3.18c)$$

$$x_c^C \in \mathbb{Z}^+, \quad \forall c, \quad (3.18d)$$

$$\sum_s C_s^{IS} x_s^S + \sum_r C_r^{IR} x_r^R + \sum_c C_c^{IC} x_c^C \leq \bar{C}^I, \quad (3.18e)$$

$$\begin{aligned} & \sum_{g \in \Psi_n^G} p_{gdt}^G(\omega) + \sum_{r \in \Psi_n^R} p_{rdt}^R(\omega) + \sum_{s \in \Psi_n^S} p_{sdt}^{SD}(\omega) \eta_s^{SD} - \sum_{\ell | s(\ell)=n} p_{\ell dt}^L(\omega) + \sum_{\ell | r(\ell)=n} p_{\ell dt}^L(\omega) \\ &= \sum_{l \in \Psi_n^L} \left(P_l^D(\omega) K_{ldt}^D - p_{ldt}^{LS}(\omega) \right) + \sum_{s \in \Psi_n^S} p_{sdt}^{SC}(\omega) + \sum_{c \in \Psi_n^C} p_{cdt}^C(\omega), \quad \forall n, \forall d, \forall t, \forall \omega \end{aligned} \quad (3.18f)$$

$$\begin{aligned} & \sum_{g \in \Psi_n^G} q_{gdt}^G(\omega) + \sum_{r \in \Psi_n^R} q_{rdt}^R(\omega) - \sum_{\ell | s(\ell)=n} q_{\ell dt}^L(\omega) + \sum_{\ell | r(\ell)=n} q_{\ell dt}^L(\omega) \\ &= \sum_{l \in \Psi_n^L} \alpha_{ldt} \left(P_l^D(\omega) K_{ldt}^D - p_{ldt}^{LS}(\omega) \right), \quad \forall n, \forall d, \forall t, \forall \omega, \end{aligned} \quad (3.18g)$$

$$V_n \leq v_{ndt}(\omega) \leq \bar{V}_n, \quad \forall n/n : \text{ref.}, \forall d, \forall t, \forall \omega, \quad (3.18h)$$

$$v_{ndt}(\omega) = V_0, \quad n : \text{ref.}, \forall d, \forall t, \forall \omega, \quad (3.18i)$$

$$v_{r(\ell)dt}(\omega) = v_{s(\ell)dt}(\omega) - \frac{\kappa_{\ell} p_{\ell dt}^L(\omega) + \chi_{\ell} q_{\ell dt}^L(\omega)}{V_0} \quad \forall \ell, \forall d, \forall t, \forall \omega, \quad (3.18j)$$

$$\underline{P}^G \leq p_{gdt}^G(\omega) \leq \bar{P}^G, \quad \forall g, \forall d, \forall t, \forall \omega \quad (3.18k)$$

$$\underline{Q}^G \leq q_{gdt}^G(\omega) \leq \bar{Q}^G, \quad \forall g, \forall d, \forall t, \forall \omega \quad (3.18l)$$

$$0 \leq p_{ldt}^{LS}(\omega) \leq P_l^D(\omega) K_{ldt}^D \quad \forall l, \forall d, \forall t, \forall \omega \quad (3.18m)$$

$$0 \leq p_{rdt}^R(\omega) \leq K_{rdt}^R \bar{P}_r^R x_r^R, \quad \forall r, \forall d, \forall t, \forall \omega \quad (3.18n)$$

$$e_{sdt}(\omega) = e_{sd,t-1}(\omega) + p_{sdt}^{SC}(\omega) \eta_s^{SC} \Delta t - p_{sdt}^{SD}(\omega) \Delta t, \quad \forall s, \forall d, \forall t, \forall \omega \quad (3.18o)$$

$$\underline{E}_s^S x_s^S \leq e_{sdt}(\omega) \leq \bar{E}_s^S x_s^S, \quad \forall s, \forall d, \forall t, \forall \omega, \quad (3.18p)$$

$$0 \leq p_{sdt}^{SC}(\omega) \leq \bar{P}_s^{SC} x_s^S, \quad \forall s, \forall d, \forall t, \forall \omega, \quad (3.18q)$$

$$0 \leq p_{sdt}^{SD}(\omega) \leq \bar{P}_s^{SD} x_s^S, \quad \forall s, \forall d, \forall t, \forall \omega, \quad (3.18r)$$

$$0 \leq p_{cdt}^C(\omega) \leq \bar{P}_c^C x_c^C, \quad \forall c, \forall d, \forall t, \forall \omega, \quad (3.18s)$$

$$0 \leq p_{cdt}^C(\omega) \leq N_{cd}^{EV}(\omega) P_{cd}^{EV} K_{cdt}^{EV}, \quad \forall c, \forall d, \forall t, \forall \omega, \quad (3.18t)$$

$$\sum_t p_{cdt}^C(\omega) \leq N_{cd}^{EV}(\omega) P_{cd}^{EVmax}, \quad \forall c, \forall d, \forall \omega, \quad (3.18u)$$

where set $\Phi^S = \{x_s^S, \forall s; x_r^R, \forall r; x_c^C, \forall c; p_{gdt}^G(\omega), \forall d, \forall t, \forall \omega; q_{gdt}^G(\omega), \forall d, \forall t, \forall \omega; p_{rdt}^R(\omega), \forall r, \forall d, \forall t, \forall \omega; q_{rdt}^R(\omega), \forall r, \forall d, \forall t, \forall \omega; p_{sdt}^{SD}(\omega), p_{sdt}^{SC}(\omega), e_{sdt}(\omega), \forall s, \forall d, \forall t, \forall \omega; p_{ldt}^L(\omega), \forall \ell, \forall d, \forall t, \forall \omega; q_{ldt}^L(\omega), \forall \ell, \forall d, \forall t, \forall \omega; p_{ldt}^{LS}(\omega), \forall l, \forall d, \forall t, \forall \omega; p_{cdt}^C(\omega), \forall c, \forall d, \forall t, \forall \omega; v_{ndt}(\omega), \forall n, \forall d, \forall t, \forall \omega\}$ includes the optimization variables of problem (3.2).

The aim of this stochastic programming problem is determining the expansion decisions that minimize the expected costs (investment and operating costs) over all scenario realizations. Scenarios are used to represent long-term uncertainties, namely the price of the electricity exchanged with the main grid, the number of EVs, and the value of peak demand, which in the ARO approach are modeled using uncertainty sets based on confidence bounds. The stochastic programming problem is solved for different numbers of input scenarios that, for the sake of clarity, are randomly generated using a normal distribution. Moreover, to be in line with the results of the ARO approach, both the “N.L.M.G.” and “L.M.G.” cases are considered. Results are provided in Table 3.13, which shows the computation time needed to solve the stochastic problem as the number of scenarios is increased. These simulations are run using CPLEX in GAMS with the same computer used for the implementation of the ARO model, namely, a MacBook Pro computer with processor Intel Core i7 at 2.7 GHz and 16 GB of RAM.

Note that the results of both methods are compared only in terms of computational performance due to the following reasons:

1. The two approaches have different objectives, namely, the stochastic programming formulation of the problem aims at determining the expansion plan that minimizes the expected costs over all scenarios, while the ARO model determines the expansion plan that minimizes the worst-case cost over the considered uncertainty set.
2. To compare the results in terms of expansion plans, it would be needed to carry out an out-of-sample assessment. To do this, future system conditions are needed to compare the cost of using the two expansions plans (those obtained from the stochastic and ARO models). One option would be to generate some out-of-sample scenarios but in such a case we would be assuming that the probability distribution of long-term uncertain parameters is known, which is a hard assumption in practice.

# Scenarios	N.L.M.G.	L.M.G.
1	0.20 s	0.26 s
10	21 s	1.13 min
20	1 min	2.27 min
50	4 min	18.18 min
80	11 min	23 min
100	16 min	No solution returned
200	Out of memory	Out of memory
1000	Out of memory	Out of memory
10000	Out of memory	Out of memory

Table 3.13: CPU times of the stochastic programming model depending on the number of scenarios

As can be observed, the number of input scenarios has a great impact on the computational performance of the stochastic model, especially when the number of considered scenarios increases and more restrictions are imposed in the model. Moreover, in order to generate accurate scenarios, information about the probability distribution of uncertain parameters is needed, which is not always available, especially for long-term uncertainties. In addition, the feasibility of the problem is only guaranteed for the considered scenarios. Thus, a large enough number of scenarios is generally needed since if the actual uncertainty realization is different to the scenarios considered, the problem may be infeasible, which may have a very negative impact on the distribution system. On the other hand, a very large number of scenarios may result in a computationally intractable problem, as highlighted by the results reported in Table 3.13. For instance, in the “L.M.G.” case with 100 scenarios, no solution is provided because the implementation is terminated by the solver and when the number of scenarios reaches 200 the computer used for simulations is no longer able to support the implementation of the model. This means that either a scenario reduction technique has to be applied or the access to a more powerful server is needed.

In this sense, the proposed ARO problem outperforms the stochastic model in two main aspects. On the one hand, the computation time needed to obtain the expansion decisions does not depend on the number of input scenarios. On the other hand, the feasibility of the problem is guaranteed provided that uncertain parameters take values within the considered confidence bounds.

3.6 Conclusions

Considering the theoretical framework of the ARO approach for the expansion planning problem of a distribution system carried out in this chapter, the general conclusions below are in order:

1. Transforming the energy system switching to renewable energy sources and developing alternative fuels are two great opportunities to reach the decarbonization targets of the European Energy Roadmap 2050. The proposed approach considers, in an integrated framework, what is needed to achieve the European decarbonization targets, namely, the planning of an electric energy system with renewable energy sources generation, storage units, network expansion, and charging stations for EVs.
2. Different sources of long-term uncertainty and short-term variability highly influence expansion decisions. The proposed approach allows an accurate modeling of them.
3. The proposed three-level formulation can be reformulated as a two-level one by merging the second-level problem and the dual of the third-level problem. The resulting problem can be efficiently solved via a column-and-constraint generation algorithm.

Regarding the case study analyzed in this chapter, the following conclusions are enumerated:

1. The expansion decisions are highly conditioned by the uncertainty budgets, especially in the “L.M.G.” cases, where the access to the main grid is limited. In particular, investments in generation and storage capacity generally increase as the values of the uncertainty budgets Λ^G and Λ^L increase, while investments in charging stations decrease. Instead, in the “N.L.M.G.” cases, the central planner only invests in charging stations since all system demand can be satisfied using the electricity withdrawn from the main grid, and there is no need to construct PV nor storage units. This happens independently of the uncertainty budget analyzed.
2. Increasing the revenues of the central planner from selling electricity to the EVs implies a raise of the number of charging stations built both under the “N.L.M.G.” and “L.M.G.” assumptions. However, in the “L.M.G.” system configuration, the central planner invests in charging stations only if these revenues are high enough.
3. Limiting the power that can be withdrawn from the main grid increases the investments in generation and storage capacity.
4. When the investment budget availability is limited, the central planner has to reduce the amount of PV power plants built and no longer invests in charging stations to contain system demand and to counteract load shedding.

5. Adopting EVs increases the electricity demand that, at the time being, is still partially satisfied by conventional and pollutant fuels, such as natural gas and coal. Anyway, the adoption of EVs allows CO₂ savings if compared with an analogous situation with lower electricity demand and diesel vehicles.
6. The proposed approach improves the results of a stochastic model in terms of computational performance.

Finally, future research includes the extension of the proposed model to a dynamic approach in which expansion decisions can be made at different points in time as proposed in García-Bertrand and Mínguez (2016), Lumbreras and Ramos (2016), and Roldán et al. (2018).

Appendix

3.A Additional data

Tables 3.A.1 and 3.A.2 report the technical data of the 69-bus distribution system and the average peak load values per node, respectively.

Line	From node	To node	$\kappa_\ell(p.u.)$	$\chi_\ell(p.u.)$
1	1	2	0.0003	0.0007
2	2	3	0.0003	0.0007
3	3	4	0.0009	0.0022
4	4	5	0.0157	0.0183
5	5	6	0.2284	0.1163
6	6	7	0.2378	0.1211
7	7	8	0.0575	0.0293
8	8	9	0.0308	0.0157
9	9	10	0.5110	0.1689
10	10	11	0.1168	0.0386
11	11	12	0.4439	0.1467
12	12	13	0.6426	0.2121
13	13	14	0.6514	0.2153
14	14	15	0.6601	0.2181
15	15	16	0.1227	0.0406
16	16	17	0.2336	0.0772
17	17	18	0.0029	0.0010
18	18	19	0.2044	0.0676
19	19	20	0.1314	0.0431
20	20	21	0.2131	0.0704
21	21	22	0.0087	0.0029
22	22	23	0.0993	0.0328
23	23	24	0.2161	0.0714
24	24	25	0.4672	0.1544
25	25	26	0.1927	0.0637
26	26	27	0.1081	0.0357
27	3	28	0.0027	0.0067
28	28	29	0.0399	0.0976
29	29	30	0.2482	0.0820
30	30	31	0.0438	0.0145
31	31	32	0.2190	0.0724
32	32	33	0.5235	0.1757
33	33	34	1.0657	0.3523
34	34	35	0.9197	0.3040
35	3	36	0.0027	0.0067
36	36	37	0.0399	0.0976
37	37	38	0.0657	0.0767
38	38	39	0.0190	0.0221
39	39	40	0.0011	0.0013
40	40	41	0.4544	0.5309
41	41	42	0.1934	0.2260
42	42	43	0.0256	0.0298
43	43	44	0.0057	0.0072
44	44	45	0.0679	0.0857
45	45	46	0.0006	0.0007
46	4	47	0.0021	0.0052
47	47	48	0.0531	0.1300
48	48	49	0.1808	0.4424
49	49	50	0.0513	0.1255
50	8	51	0.0579	0.0295
51	51	52	0.2071	0.0695
52	9	53	0.1086	0.0553
53	53	54	0.1267	0.0645
54	54	55	0.1773	0.0903
55	55	56	0.1755	0.0894
56	56	57	0.9920	0.3330
57	57	58	0.4890	0.1641
58	58	59	0.1898	0.0628
59	59	60	0.2409	0.0731
60	60	61	0.3166	0.1613
61	61	62	0.0608	0.0309
62	62	63	0.0905	0.0460
63	63	64	0.4433	0.2258
64	64	65	0.6495	0.3308
65	11	66	0.1255	0.0381
66	66	67	0.0029	0.0009
67	12	68	0.4613	0.1525
68	38	69	0.0029	0.0010

Table 3.A.1: 69-buses network data

Node	Demand (kW)
6	2.6
7	40.4
8	75
9	30
10	28
11	145
12	145
13	8
14	8
16	45.5
17	60
18	60
20	1
21	114
22	5.30
24	28
26	14
27	14
28	26
29	26
33	14
34	19.5
35	6
36	26
37	26
39	24
40	24
41	1.2
43	6
45	39.22
46	39.22
48	79
49	384.7
50	384.7
51	40.5
52	3.6
53	4.35
54	26.4
55	24
59	100
61	1244
62	32
64	227
65	59
66	18
67	18
68	28
69	28
Total	3802.19

Table 3.A.2: Average peak load values per node

Chapter 4

Analysis of long-term natural gas contracts with vine copulas in optimization portfolio problems

In this chapter we investigate the dependence risk and the optimal resource allocation of the underlying assets of a gas long-term contract (LTC) through pair-vine copulas and portfolio optimization methods with respect to different risk measures. In Europe gas is sold according to two main methods: long-term contract and hub pricing. Europe is moving towards a mix of long term and spot markets, but the eventual outcome is still unknown. The fall of the European gas demand combined with the increase of the US shale gas exports and the rise of Liquefied Natural Gas availability on international markets have led to a reduction of the European gas hub prices. On the other side, oil-indexed LTCs failed to promptly adjust their positions, implying significant losses for European gas mid-streamers that asked for a re-negotiation of their existing contracts and obtained new contracts linked also to hub spot prices. The debate over the necessity of the oil-indexed pricing is still on-going. The supporters of the gas-indexation state that nowadays the European gas industry is mature enough to adopt hub-based pricing system. With the aim of analyzing this situation and determining whether oil-indexation can still be convenient for the European gas market, we consider both spot gas prices traded at the hub and oil-based commodities as possible underlyings of the LTCs. Our results show that European LTCs will most likely remain indexed to oil-based commodities, even though a partial dependence on spot hub prices is conceded.

The model developed in this chapter have been implemented and solved in R.

4.1 Introduction

Natural gas can be sold either through long-term contracts (LTCs) or at spot price at market hubs. LTCs have been historically introduced to allow for risk sharing between gas producers and mid-streamers, which respectively face price and volume risks. These contracts have been traditionally concluded over long periods (typically 20 years or more) and are characterized by quantity and price clauses. The Take or Pay (TOP) quantity clause obligates the buyer to take a certain quantity of natural gas or to pay for it. The Price Indexation (PI) clause relates the price at which gas is bought to some index on the market that has been traditionally represented by the price of crude and oil products. The price clause provides producers with some price stability and reduction of revenue volatility, which are indispensable for ensuring investments in new infrastructures which are very expensive. Moreover, the quantity and price clauses also allow for hedging the mid-streamers' volume risk (see Abada et al. (2017)).

The hub pricing approach was firstly introduced in the nineties in the US and in the UK and it is now developing in Europe. In this system, natural gas is traded, every day, on a spot market that determines prices and volume on the short term. International natural gas market is organized in different ways depending on the considered areas. North America is essentially organized on the basis of Henry Hub spot market. The price on this market is currently very low, at about 3\$/MBtu,¹ because of the development of unconventional (shale) gas. The Henry Hub is the best-known of all natural gas trading points. It is both a physical distribution hub for pipeline gas and a pricing point, since it is the basis of spot market trading and of futures trading on the New York Mercantile Exchange (NYMEX).

On the other side, Asia is mainly supplied by Liquefied Natural Gas (LNG) that is traded through expensive LTCs. The Asian LNG long-term contracts has been set on the basis of the average of Japanese customs-cleared crude oil price that is the Japan Crude Cocktail (JCC). After the Fukushima disaster, Japan has significantly increased its imports of LNG, mainly supplied by Australia, Malaysia and Qatar through LTCs (see International Group of Liquefied Natural Gas Importers (2018)).

Spot deals exist, but they are bilateral as, up to now, there is no hub in Asia in contrast with America and Europe. Some experts advocate a move from the current system of LTCs to hub-pricing system (e.g International Energy Agency (2013)). Singapore has proposed itself as a possible gas hub for Asia as well as Tokyo and Shanghai (see Xunpeng (2016)).

Except for the UK, Europe is still dominated by LTCs, though spot markets are growing and are

¹See <http://www.eia.gov/dnav/ng/hist/rngwhhdd.htm>

expected to develop further. In the UK, gas is largely traded at the National Balancing Point (NBP) spot market. NBP is in operation since the late 1990s and is the longest-established spot-traded natural gas market in Europe. It is characterized by high liquidity² and the resulting spot price is widely used as an indicator for European wholesale gas market. In continental Europe, Zeebrugge (ZEE) and the Title Transfer Facility (TTF), respectively located in Belgium and in the Netherlands, are the two dominant spot market places and many others are emerging (see Melling (2010)).

The coexistence of LTCs and hub-pricing systems on the European market implies that the natural gas is traded at two different prices on the same market. Depending on the conditions, spot price can be higher or lower than long-term contractual gas, leading to possibly difficult situations for companies loaded with high TOP gas price against the low spot prices. This is what happened in the last years in Europe, where the combined effects of the increase of the US shale gas exports, the reduction of European gas demand due to the economic crisis, and the increased availability of uncommitted LNG from Qatar led to a new supply/demand balance that was reflected into low gas prices at the European hubs. On the other side, oil-indexed LTCs failed to promptly adjust their positions implying significant losses for European gas mid-streamers what were committed by the TOP clause to buy quantities of gas higher than those required at higher prices. As a consequence, European mid-streamers have re-negotiated the LTCs to make them more flexible and closer to spot gas prices. These re-negotiations have resulted into a decline of oil-indexation and hub-linked pricing has rapidly become the basis for an increasing number of transactions in the UK and in the Northern Western Europe (see Franza (2014); Yafimava (2014)). As indicated by Chyong (2015), leading gas suppliers, such as Statoil, GasTerra, Sonatrach and Gazprom, have been forced to modify their LTC prices and volumes in Europe. However, these gas suppliers have assumed different attitudes: Gazprom and Sonatrach has defended oil indexation and has offered retroactive discounts on existing contracts by introducing either limited degree of spot indexation or reduced minimum TOP provisions. On the other side, GasTerra and Statoil have behaved in a more flexible way, by conceding more spot indexation in their contracts. Notwithstanding this, oil is still accounting for the 95% of price formation within the Continental Europe against the 30% of price formation in the UK (see Theisen (2014)). This shows that oil-indexed contract prices still exercise a strong influence over gas prices in Europe. The reliability of prices set on spot markets is one of the main reasons used by the opponents of the spot indexation for LTCs (see Frisch (2010)). Liquidity, transparency, and the ability to attract a significant number of market players are necessary

²The liquidity of a gas hub can be defined as the ratio between the total volume of trade on the hub and the volume of gas consumed in the area served by the hub.

for a hub to become a price maker (see Heather (2012)). For the time being, the NBP and TTF are the only two hubs in Europe with sufficient liquidity (see Heather and Petrovich (2017)). However, some experts advocate a move from the current system of long-term contracts to hub pricing (see Stern and Rogers (2014)). Even though their arguments make sense, these are not supported by numerical analysis.

From these evidences, it turns out that the role of the European spot gas markets and their impacts on LTCs are becoming extremely challenging issues. This work sheds some light on the LTCs re-negotiation, taking into account the mid-streamers' requests. An integrated framework that combines vine copula and optimization techniques traditionally used in the contest of portfolio management. The adopted methodology is well studied in the literature and it is considered mature enough for this task. This allows us to obtain reliable results that can provide usable information and could be easily reproducible by gas market players and practitioners. Our analysis is structured as follows: we estimate, via Pair-Copula Constructions (PCC), the dependence risk structure across the underlyings of LTCs on natural gas. In order to reflect the above mentioned European hybrid pricing system, based on the symbiotic coexistence of oil-indexed contracts and gas-indexed hub prices, we consider the prices of oil-based products and natural gas traded at the hub as component of the gas pricing formula. (see Section 4.3 for more details). We then combine vine copula models and classical portfolio optimization methods to construct the optimal underlying portfolio, applying different performance measures. Copula models represent a suitable tool to this scope. In particular, copula is a function that combines marginal distributions to form multivariate distributions. The application of copulas is very popular in several fields, like finance, insurance, financial economics and econometrics (see e.g. Durante and Sempi (2015); Genest and Bourdeau-Brien (2009); Krzemienowski and Szymczyk (2016); Li and You (2014); Saida and Prigent (2018); Tran et al. (2017)). Nowadays, the modeling of stochastic dependence via copulas has led to an increasing attention also in the commodity market (see e.g. Accioly and Aiube (2008); Aloui et al. (2013); Bassetti et al. (2018); Grégoire et al. (2008); Jäschke (2014); Lu et al. (2014); Reboredo (2011); Wen et al. (2012)).

While there is a wide range of possible alternative copula functions for the bivariate case, in the multivariate setting the use of families different from normal and student-t is rather scarce, due to computational and theoretical limitations (see e.g. Joe (1997); Nelsen (1999)). For this reason, in order to represent a multivariate copula with suitable sets of bivariate copulas, Joe (1996) introduced the PCC approach, later discussed in detail by Aas et al. (2009); Bedford and Cooke (2001, 2002); Kurowicka and Cooke (2006). A collection of potentially different bivariate copulas is used to con-

struct the joint distribution of interest via PCCs, allowing to represent different types and strengths of dependence in an easy way. PCCs constitute a flexible and very appealing tool for financial analysis, (see Brechmann and Czado (2013); Dalla Valle et al. (2016); Dißmann et al. (2013)). The vine copula models considered for the analysis of the portfolio's dependence risk overcome the restrictive and deterministic features of the bivariate copulas and traditional measures of correlation, due to their suitability in capturing the non-normality, tail dependence and volatility clustering of assets returns. Recently Hernandez (2014) show how the vine copula approach can be appropriately used to investigate the dependence structure among the different components of energy portfolio as well as to derive implications for portfolio risk management. In this work, we investigate via PCC, the dependence risk structure across the underlying assets of LTCs on natural gas. We define the optimal portfolio composition under different performance measures.

To the best of our knowledge the optimal composition of the assets, commonly used to price the gas LTCs, has not been investigated using the aforementioned approach. In doing this, we consider both the traditional oil-based commodities and spot gas prices to address the debate over oil/spot indexation related to the re-negotiation of European LTCs. Our results confirm the effectiveness of the hybrid pricing system currently existing in the continent, but indicates that oil should still play an important role in the definition of the price of the LTCs.

The remainder of this work is organized as follows: Section 4.2 briefly presents the PCC and the vine copulas; Section 4.3 provides the data analysis and introduces the pair vine copula that we use in our study; Section 4.4 overviews the optimization portfolio problems that we solve to estimate the risk of the optimal portfolio of the underlying assets of LTCs, and illustrates the results of our analysis. Finally, concluding remarks are given in Section 4.5.

4.2 PCC approach and vine copulas

PCC is a multivariate copula constructed by using only bivariate copula or pair-copulas as building blocks. All copulas involved in the decomposition may be selected freely among the wide range of bivariate copula family that are capable of modeling joint distribution with different characteristics. Hence, PCC allows high flexibility in representing complex structures of dependence among multivariate data. It is based on the decomposition of a d -dimensional joint density function $f(x_1, \dots, x_d)$ of the random vector $X = (X_1, \dots, X_d)$, as a product of conditional densities:

$$f(x_1, \dots, x_d) = f_d(x_d) \times f_{d-1|d}(x_{d-1} | x_d) \times \dots \times f_{1|2\dots d}(x_1 | x_2 \dots, x_d) . \quad (4.1)$$

Each term in (4.1) can be decomposed using Sklar theorem (see Sklar (1959)) to express the conditional density for a generic element x_j conditioned on the d -dimensional vector $\mathbf{v}$ as in (4.2):

$$f_{x_j|\mathbf{v}}(x_j|\mathbf{v}) = c_{x_j,v_l|\mathbf{v}_{-l}}(F_{x_j|\mathbf{v}_{-l}}(x_j|\mathbf{v}_{-l}), F_{v_l|\mathbf{v}_{-l}}(v_l|\mathbf{v}_{-l})) \times f_{x_j|v_{-l}}(x_j|\mathbf{v}_{-l}), \quad (4.2)$$

where v_l is an arbitrary component of $\mathbf{v}$, $\mathbf{v}_{-l}$ denotes the $(d-1)$ dimensional vector without v_l , $c_{x_j,v_l|\mathbf{v}_{-l}}(\cdot, \cdot)$ is the conditional pair copula density and $F_{x_j|\mathbf{v}_{-l}}(\cdot|\cdot)$ is the conditional distribution of x_j given $\mathbf{v}_{-l}$. More precisely, for every j , Joe (1996) proves that:

$$F_{x_j|\mathbf{v}}(x_j|\mathbf{v}) = \frac{\partial C_{x_j,v_l|\mathbf{v}_{-l}}(F_{x_j|\mathbf{v}_{-l}}(x_j|\mathbf{v}_{-l}), F_{v_l|\mathbf{v}_{-l}}(v_l|\mathbf{v}_{-l})|\theta)}{\partial F_{v_l|\mathbf{v}_{-l}}(v_l|\mathbf{v}_{-l})}, \quad (4.3)$$

where the bivariate copula function is specified by $C_{x_j,v_l|\mathbf{v}_{-l}}$, with parameters θ . In working with copula models a function $h(x, v, \Theta)$ can be defined in order to represent the conditional distribution function when x and v are uniform, i.e. $f(x) = f(v) = 1$, $F(x) = x$ and $F(v) = v$. The h-function can be calculated as follows:

$$h(x, v, \Theta) = F(x|v) = \frac{\partial C_{x,v}(x, v, \Theta)}{\partial v}, \quad (4.4)$$

where the second parameter of $h(x, v, \Theta)$ always corresponds to the conditioning variable and Θ denotes the set of parameters for the copula of the joint distribution function of x and v . For high-dimensional distributions the number of possible pair-copulas constructions is manifold. In order to organize them, Bedford and Cooke (2001, 2002) have introduced a graphical model denoted regular vines that depict multivariate copulas built up using a cascade of bivariate copulas (or pair-copulas). This allows to understand which conditional specifications are used to describe the joint distribution. A regular vine $V(d)$ on d variables is a nested set of trees T_i where $i = 1, \dots, d-1$. In particular the first tree has d nodes and $d-1$ edges that represents the pair-copula densities between the nodes. While, the j trees have $d+1-j$ nodes deriving from the edges of tree $j-1$ and $d-j$ edges that are the conditional pair-copula densities. Moreover the proximity condition states that if the nodes of tree $j+1$ are connected by an edge, than the corresponding edges in tree j share a common node. According to Kurowicka and Cooke (2006) the joint of a random vector $X = (X_1, \dots, X_d)$ following an R -vine distribution can be written as:

$$f(x_1, \dots, x_d) = \left[\prod_{k=1}^d f_k(x_k) \right] \cdot \left[\prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(x_{j(e)}|x_{D(e)}), F(x_{k(e)}|x_{D(e)})) \right]. \quad (4.5)$$

with node set $\mathcal{N} := \{N_1, \dots, N_{d-1}\}$ and edge set $\mathcal{E} := \{E_1, \dots, E_{d-1}\}$. Each parameter $e = j(e), k(e) | D(e)$ is an edge, while $c_{j(e), k(e) | D(e)}$ represents a bivariate conditional density copula. $j(e)$ and $k(e)$ are the conditioned nodes and $D(e)$ is the conditioning set. The union $\{j(e), k(e), D(e)\}$ is called constraint set. The right term of equation (4.5) which involves $d(d-1)/2$ bivariate copula densities, is called an R -vine copula. Special classes of R -vines are Canonical vines (C -vines) and Drawable vines (D -vines). A C -vine is a regular vine where each tree T_i has a unique node that is connected to $d-i$ edges; while a D -vine is a regular vine where each node is connected to no more than two other nodes. Each tree in a C -vine is a star with one unique node that connects to all other nodes, whereas a D -vine is represented by line trees. In a C -vine, at the first root node at level 1 of the nested set of trees, the key variable presents the highest correlation value in regard to the other variables and governs the dependence structure among the others. Intuitively, we use a C -vine to describe a scenario where one variable dominates the others, whereas in the D -vines we do not assume the existence of a particular node dominating the dependencies. More precisely, the joint density function $f(x_1, \dots, x_n)$ of a C -vine of dimension d takes the following form:

$$\begin{aligned} f(x_1, \dots, x_d) &= \\ &= \prod_{k=1}^d f_k(x_k) \cdot \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{i, i+j | 1, \dots, i-1} (F(x_i | x_1, \dots, x_{i-1}), F(x_{i+j} | x_1, \dots, x_{i-1}) | \theta_{i, i+j | 1, \dots, i-1}). \end{aligned} \quad (4.6)$$

In a similar way, the joint density function of a D -vine is given by:

$$\begin{aligned} f(x_1, \dots, x_d) &= \\ &= \prod_{k=1}^d f_k(x_k) \cdot \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{j, j+i | j+1, \dots, j+i-1} (F(x_j | x_{j+1}, \dots, x_{j+i-1}), F(x_{j+i} | x_{j+1}, \dots, x_{j+i-1}) | \theta_{j, j+i | j+1, \dots, j+i-1}). \end{aligned} \quad (4.7)$$

Differently from C -vines and D -vine the joint density function of the R -vine can vary significantly according to the statistical feature of the multivariate distribution being modeled. For this reason, Morales-Napoles (2010); Dißmann et al. (2013) proposed an efficient method for storing the indices of the pair-copula. It relies on the specification of a lower triangular matrix $M = (m_{i,j} | i, j = 1, \dots, d) \in \{0, \dots, d\}^{d \times d}$, whose diagonal entries $m_{i,i}$ are the nodes $1, \dots, d$ of the first tree. Each row

from the bottom up represents a tree. The conditioned sets of a node are determined by a diagonal entry and the corresponding column entry of the row under consideration, while the the column entries below this row provides the conditioning set. Corresponding copula types can also be stored in matrices similar to M (see Section 4.3.2 for more details).

4.3 Data and dependence structure

4.3.1 Data analysis

With the aim of analyzing the re-negotiation of the European LTCs, we consider both the oil-based commodities traditionally used to determine the LTCs price and the main spot gas prices traded at the hub.

In particular the historical daily prices of the following assets have been taken into account: Crude Ice Brent DTD, Gasoil NWECIF, Jet Fuel NWECIF, Naphtha NWECIF, Lsfo 1% NWECIF, Gas NBP 1stMonth, Gas HenryHub 1stMonth. The label “NWE” stands for the reference market “North West Europe”, while “CIF” indicates the “Cost, Insurance and Freight” that are the costs included in the prices. For the sake of simplicity, in the rest of the work, we denote the seven time series as follows: “Brent”, “Gasoil”, “JetF”, “Naphtha”, “Lsfo”, “Gas NBP” and “Gas HenryHub”.

We analyze the period from the 4th of January 2012 to the 24th of July 2014 for a total of 647 observations. The first five time series, referred to oil and its by-products, are spot prices. The last two series, respectively referred to the natural gas traded at the NBP and Henry Hub spot markets, are the first month future prices that can be considered as good proxy of the gas spot prices. Data are provided by Datastream Thomson Reuters and all the numerical computations are run in R 3.3.2.

For each historical daily price time series, we construct the log returns and we report their basic statistics in Table 4.1. There is evidence of negative skewness and significant excess kurtosis.

Fig 4.1 shows the log returns, the 30-days horizon rolling standard deviation on log returns, and the volatility associated with the Gas NBP time series. In particular, figure 4.1b shows the volatility clustering, namely “*large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes*”, as observed by Mandelbrot (1963) (for the other series, see Fig. 4.A.1-4.A.6 in Appendix 4.5).

The volatility clustering is confirmed also by the sample autocorrelation function (ACF) of the squared mean adjusted log returns (see left hand side of Fig. 4.2 for the GAS NBP series and Fig.

Table 4.1: Basic statistics of log return time series referred to the period January 4, 2012 - July 24, 2014

	Mean	Max	Min	Std. Dev	Skewness	Kurtosis
Brent	-0.01%	6.57%	-5.67%	1.26%	-0.068	2.063
Gasoil	-0.01%	5.08%	-3.11%	1.06%	0.049	1.077
JetF	-0.01%	5.05%	-3.98%	1.03%	0.075	1.599
Naphtha	0.00%	5.27%	-7.72%	1.30%	-0.440	3.654
Lsfo	-0.02%	3.54%	-6.80%	1.13%	-0.758	4.551
Gas NBP	-0.04%	9.10%	-7.69%	1.59%	0.294	3.620
Gas HenryHub	0.04%	13.27%	-11.93%	2.78%	0.290	2.342

4.A.7-4.A.8 in Appendix 4.5 for other series).

In addition, we observe that all log return series are stationary with respect to the ADF (Dickey and Fuller (1981)) and the PP tests (Phillips and Perron (1988)). The series are also stationary compared to the KPSS test (Kwiatkowski et al. (1992)) with the exception of the Gas NBP log return series. The KPSS is negligible for the latter since only the ACF plot of the differentiated series shows the presence of unit root.

Table 4.2: Unit root test results for the considered series

P-Value	Brent	Gasoil	JetF	Naphtha	Lsfo	Gas NBP	Gas HenryHub
ADF	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
PP	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
KPSS	> 0.1	> 0.1	> 0.1	> 0.1	> 0.1	0.013	> 0.1

For modeling the volatility of the series, we test several GARCH models and, for each one, we select the best model according to the significance of parameters, the log-likelihood value and the information criteria. In the case of extreme market events, dummy variables associated with these events have been used to model volatility spikes. Every series is filtered using a TGARCH model with a skewed-t distribution for innovations that allows to capture the asymmetry in volatility (i.e. the leverage effect). ARMA models have been used to compensate for autocorrelation, modeling the conditional mean where needed. Table 4.3 reports the results of the ARMA-GARCH fitting procedure. The application of the Ljung-Box test on both standardized residuals standardized squared residuals shows the absence of series correlation and the ARCH-LM test (see Fisher and O. Gallagher (2012)) confirms the adequacy of the ARCH processes. The ACF and the Partial ACF (PACF) of both the log return series and the corresponding residuals are reported in Fig 4.3-4.4 for the GAS

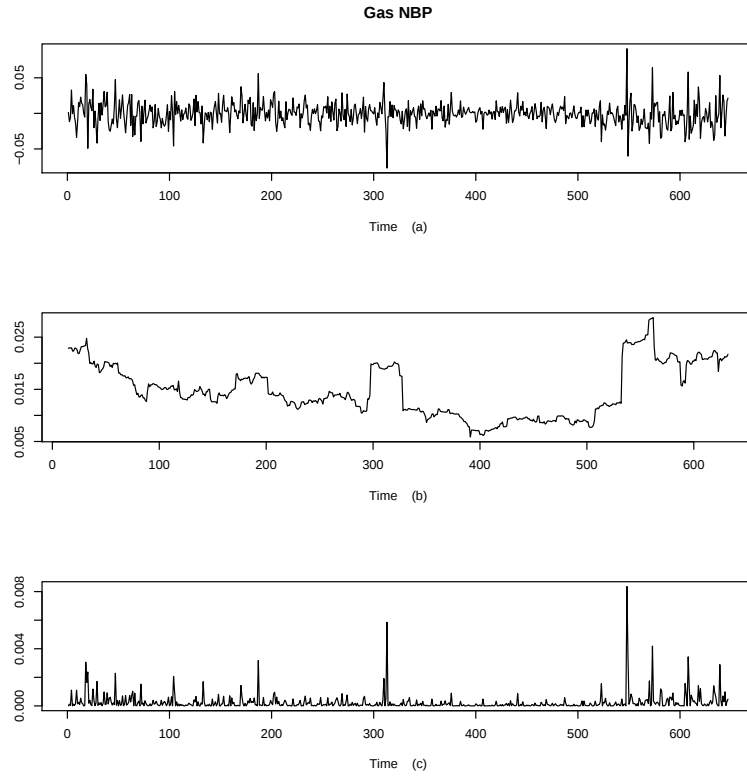


Figure 4.1: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Gas NBP time series

NBP series and Fig. 4.A.9-4.A.12 in Appendix 4.5 for other series. This indicates that there is no evidence of autocorrelation at any lag.

The reduction of the volatility after fitting the ARMA-GARCH models is confirmed by the ACF of the squared mean adjusted residuals reported in the right hand side of Fig. 4.2 for the GAS NBP series and Fig. 4.A.7-4.A.8 in Appendix 4.5 for other series.

4.3.2 Pair-Copula Constructions

Using the probability integral transformation on the standardized residuals we obtain copula u-data. In order to avoid the misspecification of the margins that may lead to the bias of the copula parameter estimates, we perform the Berkowitz test on the u-data. The outcome of this test does not indicate evidence against the uniform (0,1).

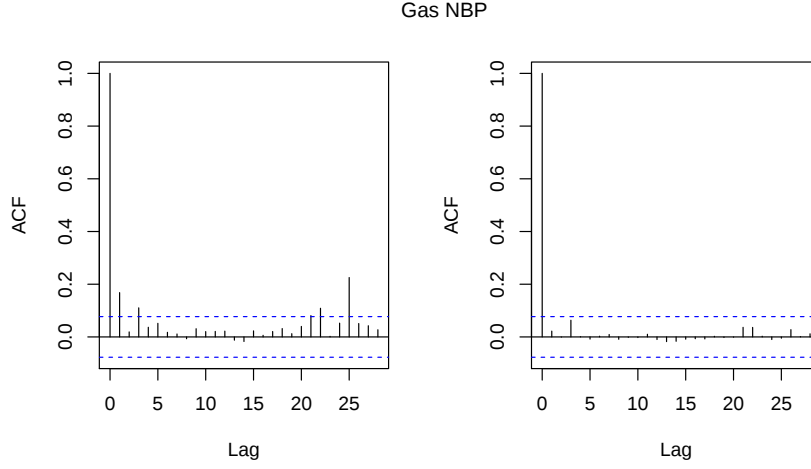


Figure 4.2: ACF of the squared mean adjusted log return series and ACF of the squared mean adjusted residuals of Gas NBP log return series.

Table 4.4 reports the pair-wise Kendall's τ correlation between the series. This allow us to construct the vine copula models using the maximum spanning tree algorithm proposed by Czado et al. (2012). We can note that if the strength of dependence is rather small, a good start of a bivariate data analysis is an independent test based on Kendall's τ correlation measure (see Genest and Favre (2007)). In the following, we estimate and analyze three different vine structures, namely C -vine, D -vine, R -vine together with the multivariate Student's t and the multivariate Gaussian copulas. Note that the multivariate Gaussian and multivariate Student's t copulas are both simplified PCCs and can be represented as an R -vine copula. In particular, the former is an R -vine with Gaussian pair-copulas, where the parameters are determined by the associated partial correlation; whereas the multivariate Student's t copula can be represented through an R -vine with bivariate t -copulas as building blocks under specific assumptions on the parameters (see Czado et al. (2009)). The R packages by Hofert et al. (2017); Schepsmeier et al. (2018) have been used for the analysis conducted in this section.

We first apply the C -vine copula model to account for the dependence structure. The application of the Kendall's τ independence test, with a confidence level of 5%, underlines the independence of the pair Gasoil-Gas HenryHub series as shown in the first C -vine tree of Figure 2. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) tests are used to select copulas. First, all available copulas are fitted with Maximum Likelihood Estimation (MLE). Then, the criteria are computed for all available copula families and the family with the minimum value is chosen.

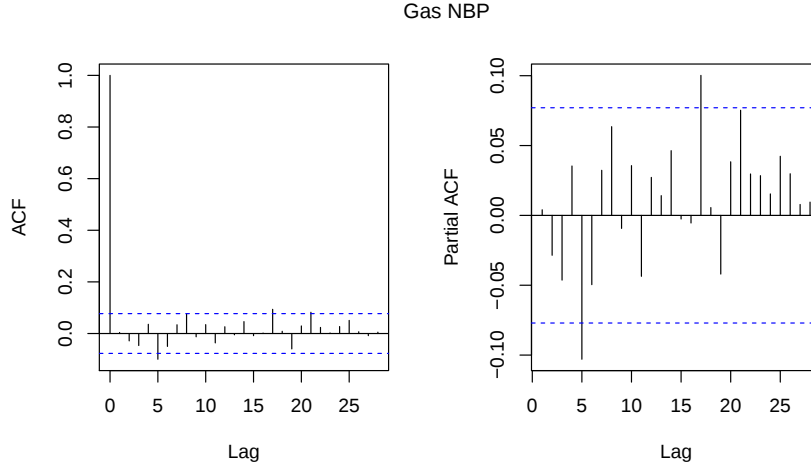


Figure 4.3: ACF and PACF of Gas NBP log return series

The parameters of the PCC can be evaluated using any multivariate copula estimator such as the maximum pseudo likelihood (MPL) estimator or the inference function for margins (IFM). However the computational effort increases exponentially with the dimension. Therefore the sequential method proposed by Aas et al. (2009), is used to estimate the parameters level by level. The selected copulas in the first tree are estimated using the MLE method. To calculate the observations (i.e. conditional distributions functions) of the second C -vine tree we use the h -functions (see (4.4) in Section 4.2). This procedure is iterated tree by tree. After having fitted all the trees, a joint MLE is provided in order to improve the estimation. This procedure requires the observations to be independent over time, in fact the PCC has been fitted on the standardized residuals obtained by filtering the original series with the ARMA-TGARCH models previously described.

In a straightforward way, we can represent the C -vine copula density factorization using the specification matrix M , while the T matrix is the copula type matrix where each row corresponds to a specific tree and each number denotes the type of pair-copula family.³

³Copula family type: 0 = Independence copula; 1 = Gaussian copula; 2 = Student-t copula (t-copula); 3 = Clayton copula; 5 = Frank copula; 13 = rotated Clayton copula (180 degrees); 14 = rotated Gumbel copula (180 degrees); 16 = Rotated Joe copula (180 degrees);

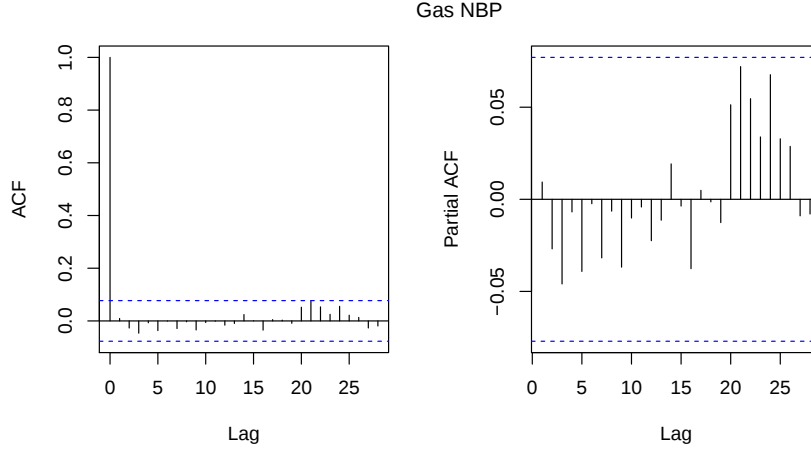


Figure 4.4: ACF and PACF of Gas NBP residuals

$$M = \begin{pmatrix} 7 \\ 6 & 6 \\ 5 & 5 & 5 \\ 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$T = \begin{pmatrix} 0 \\ 0 & 5 \\ 5 & 5 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 13 & 2 & 2 \\ 0 & 3 & 1 & 14 & 2 & 2 \end{pmatrix}$$

We recall that Gaussian and Student-t copulas accounts for a symmetric tail dependence. In particular, Gaussian copula is designed to model the dependence in the center of the joint distributions, while the Student-t copula is both lower and upper-tail dependent. Student-t copula allows for joint

Table 4.3: p-values of ARMA GARCH models. The corresponding statistics are reported in parentheses. WLB=Weighted Ljung-Box Test, WALM= Weighted Arch LM Test. S.R= Standardized Residuals. S.S.R= Standardized Squared Residuals

Series	Brent	Gasoil	JetF	Naphtha	Lsfo	Gas NBP	Gas HenryHub	
GARCH Model	TGARCH(1,2)	TGARCH(1,1)	TGARCH(2,0)	TGARCH(1,2)	TGARCH(1,2)	TGARCH(1,1)	TGARCH(1,1)	
ARMA Model	ARFIMA(1,0,1)	ARFIMA(0,0,0)	ARFIMA(0,0,0)	ARFIMA(9,0,0)	ARFIMA(14,0,0)	ARFIMA(17,0,0)	ARFIMA(13,0,0)	
WLB Test on S.R.	Lag[1]	0.9995 (4.364e-07)	0.9401 (0.005655)	0.7433 (0.1073)	0.9692 (0.001495)	0.8531 (0.03427)	0.8179 (0.05303)	0.7959 (0.0669)
	Lag[2*(p+q)+(p+q)-1]	1 (1.093)	0.9908 (0.008741)	0.9089 (0.1175)	1 (6.642716)	1 (10.24734)	1 (18.76113)	1 (7.8402)
	Lag[4*(p+q)+(p+q)-1]	0.6123 (4.330)	0.83 (1.1283)	0.533 (2.3745)	0.9993 (12.9167)	0.9974 (24.0307)	0.7594 (39.0382)	1 (16.8326)
	Lag[1]	0.2416 (1.371)	0.5683 (0.3256)	0.8934 (0.01796)	0.8808 (0.02248)	0.6537 (0.2013)	0.2902 (1.119)	0.3185 (0.995)
WLB Test on S.S.R.	Lag[2*(p+q)+(p+q)-1]	0.1808 (6.604)	0.9592 (0.4904)	0.9904 (0.22823)	0.6261 (3.32382)	0.286 (5.5291)	0.5591 (2.257)	0.225 (4.245)
	Lag[4*(p+q)+(p+q)-1]	0.2619 (9.213)	0.9871 (0.9792)	0.8128 (2.66476)	0.4987 (7.05544)	0.3085 (8.7125)	0.7089 (3.298)	0.3887 (5.285)
	ARCH Lag[4]	0.1355 (2.228)	0.8308 (0.04564)	0.7293 (0.1198)	0.3062 (1.047)	0.1477 (2.096)	0.1866 (1.745)	0.5247 (0.4047)
WALM Tests	ARCH Lag[6]	0.3469 (2.750)	0.9045 (0.42836)	0.9347 (0.3207)	0.5359 (1.802)	0.1811 (4.078)	0.5012 (1.868)	0.8632 (0.5685)
	ARCH Lag[8]	0.4318 (3.637)	0.9521 (0.73775)	0.856 (1.3236)	0.6851 (2.279)	0.2384 (5.100)	0.7118 (2.026)	0.9105 (1.0185)

Table 4.4: Kendall's τ correlation between u-data

	Brent.u	Gasoil.u	JetF.u	Naphtha.u	Lsfo.u	Gas NBP.u	Gas HenryHub.u
Brent.u	1.0000	0.6037	0.6002	0.3143	0.2750	0.0851	0.0213
Gasoil.u	0.6037	1.0000	0.7805	0.2887	0.2608	0.0847	0.0341
JetF.u	0.6002	0.7805	1.0000	0.2797	0.2524	0.0863	0.0408
Naphtha.u	0.3143	0.2887	0.2797	1.0000	0.2976	0.1004	-0.0428
Lsfo.u	0.2750	0.2608	0.2524	0.2976	1.0000	0.0990	-0.0025
Gas NBP.u	0.0851	0.0847	0.0863	0.1004	0.0990	1.0000	0.0111
Gas HenryHub.u	0.0213	0.0341	0.0408	-0.0428	-0.0025	0.0111	1.0000

extreme events either in both bivariate tails or none of them. Moreover, Frank copula models the dependence in the centre of distribution, i.e. strong dependence in non-extreme scenarios. With respect to the bivariate Gaussian, Frank copula can account for the non linearity in the center of the joint distribution, while the first one only focuses on linear dependence relationship. Finally, Clayton copula is only lower-tail dependent and is characterized by the asymmetric correlation. Together with the 180° rotated Joe and Gumbel is adequate to model greater dependence in the negative tail. The 180° rotated Clayton is suitable to model dependence in the positive tail.

The D -vine structure is reported in Fig. 4.6. The results of the independence test based on the Kendall's τ correlation measure with a 5% confidence level points out the independence of the pair Gas NBP-Gas HenryHub as shown in the first tree of the D -vine structure in figure 3. In the following the specification matrix M and copula type matrix T are reported for the D -vine

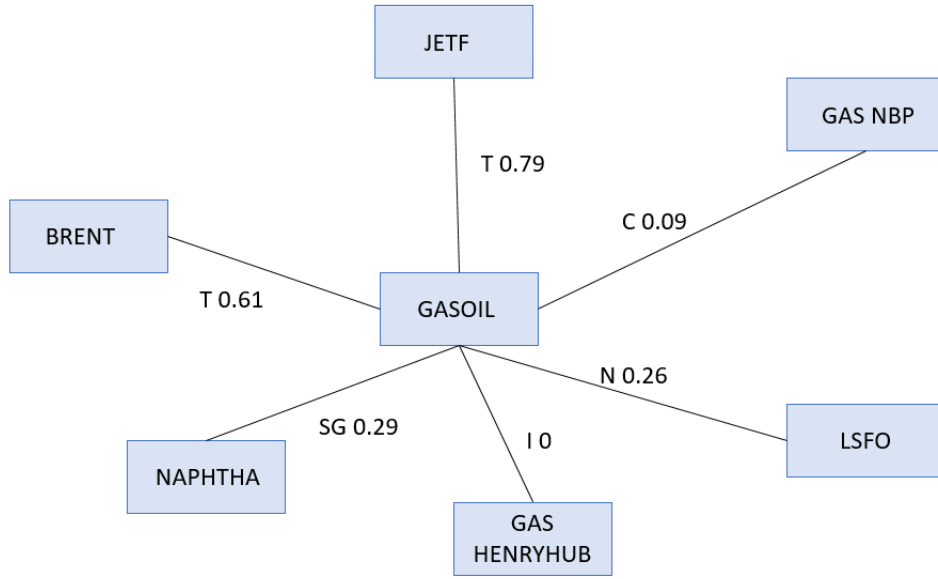


Figure 4.5: First tree of the C -vine with a 5% confidence level. The letters reported between the root nodes indicate the type of the bivariate copulas used to model the dependence, while the numbers refer to the corresponding Kendall's τ correlation.

structure.

$$M = \begin{pmatrix} 7 \\ 1 & 6 \\ 2 & 1 & 5 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 & 3 \\ 5 & 4 & 3 & 2 & 1 & 2 \\ 6 & 5 & 4 & 3 & 2 & 1 & 1 \end{pmatrix}$$

$$T = \begin{pmatrix} 3 \\ 0 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 5 & 5 & 2 \\ 0 & 1 & 2 & 1 & 2 & 2 \end{pmatrix}$$

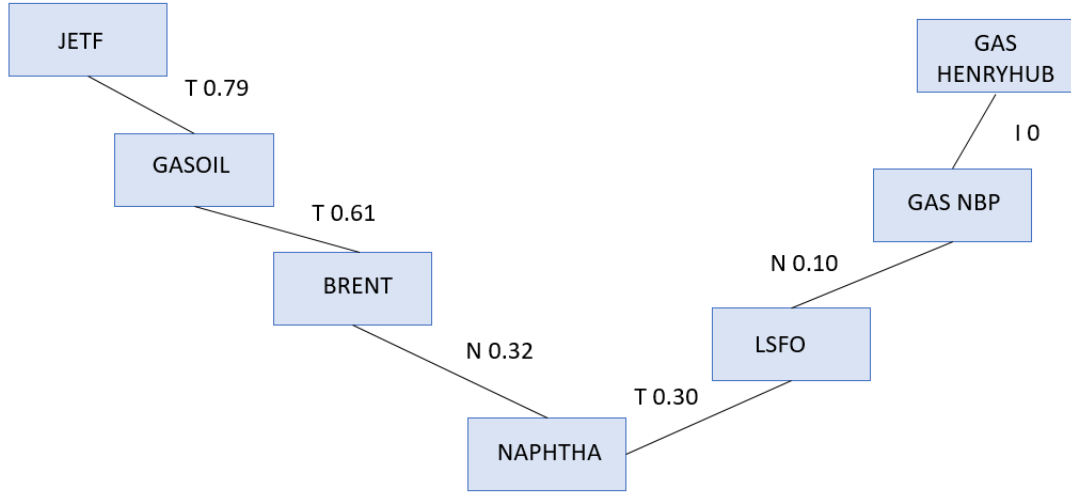


Figure 4.6: First tree of the D -vine with a 5% confidence level. The letters reported between the root nodes indicate the type of the bivariate copulas used to model the dependence, while the numbers refer to the corresponding Kendall's τ correlation.

The R -vine structure is reported in Fig. 4.7 and the respective specification matrix M and copula type matrix T are as follows.

$$M = \begin{pmatrix} 7 & & & & & & \\ 6 & 1 & & & & & \\ 5 & 6 & 2 & & & & \\ 1 & 5 & 6 & 3 & & & \\ 2 & 4 & 5 & 6 & 5 & & \\ 3 & 3 & 4 & 5 & 6 & 4 & \\ 4 & 2 & 3 & 4 & 4 & 6 & 6 \end{pmatrix}$$

$$T = \begin{pmatrix} 0 & & & & & & \\ 0 & 0 & & & & & \\ 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & 0 & & & \\ 1 & 1 & 16 & 2 & 0 & & \\ 1 & 2 & 1 & 2 & 2 & 0 & \end{pmatrix}$$

The multivariate Gaussian copula is often used for modeling multivariate data, assuming linear

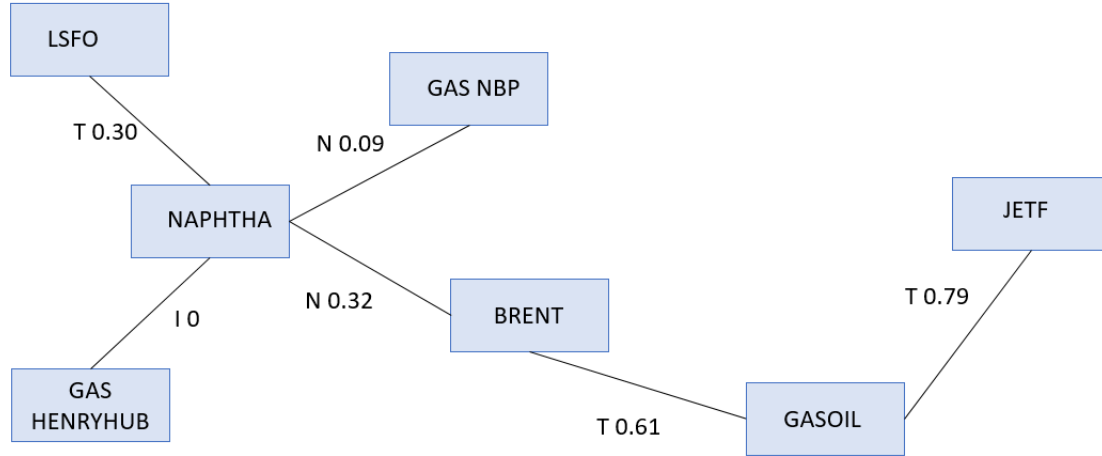


Figure 4.7: First tree of the R -vine with a 5% confidence level. The letters reported between the root nodes indicate the type of the bivariate copulas used to model the dependence, while the numbers refer to the corresponding Kendall's τ correlation.

dependence structure and no tail dependencies. The multivariate Student's t copula also assumes linear dependence, but accounts for tail dependence with the restriction of same tail heaviness within the distribution. We fit different multivariate Student's t copula by fixing 5, 10, 15 and 20 degrees of freedom (df). T-copulas families with 10, 15, and 20 df are not rejected at a 5% significance level. In the following, we consider the statistics of the 10 df Student's t copula whose p-value resulting from the multiplier goodness-of-fit (gof) test is equal to 28%.

Log-likelihood, AIC and BIC are statistics commonly used in the literature to compare different copula structures and determine their relative ranking (see, e.g., Schepsmeier (2016); Schepsmeier and Czado (2016)). The value of the log-likelihood of the estimated copula models is a gof measure, while AIC and BIC criteria are other comparison measures that take into account model complexity. The computation of these statistics for our models allows us to establish a preference for vine structures with respect to multivariate Gaussian and Student's t copulas (see Table 4.5). In particular among the vines, C -vine has the highest AIC, while R -vine is preferred in terms of BIC due to its lower number of parameters.

The comparison of the three T matrices of the considered vine structures provides additional information on tail dependence. More specifically, about 5-30% of the selected pair-copulas have either lower or upper tail dependence, i.e Clayton, Gumbel and Joe, while 20-25% have both upper and lower tail dependence modeled with a Student's t copula. Furthermore, tail dependence is non-negligible since the tail dependence coefficient calculated from the estimated pair-copula is

Table 4.5: Log-likelihood, AIC and BIC for the estimated copula models

	<i>C</i> -vine	<i>D</i> -vine	<i>R</i> -vine	Student's t	Gaussian
Log-likelihood	1336	1321	1328	1218	1158
AIC	-2632	-2607	-2626	-2394	-2300
BIC	-2547	-2531	-2563	-2300	-2265

on average greater than 0.2.

In conclusion, vine-copula structures are preferred according to the results of Table 4.5. Moreover, thanks to their flexibility, vine-copula structures can account for asymmetric estimated tail dependence. For all these reasons, we exclude the multivariate Gaussian and Student's t copulas from further analysis. As an additional step, we run the ECP and ECP2 gof tests on the three selected vine structures. These are non-parametric tests based on the Cramer-von Mises (CvM) and Kolmogorov-Smirnov (KS) test statistics (see Schepsmeier (2016) for more details on the tests).

Table 4.6: Goodness-of-fit test on *C*-vine, *D*-vine and *R*-vine with bootstrap repetition rate $x = 200$.

	ECP (CvM)	ECP (KS)	ECP2 (CvM)	ECP2 (KS)
<i>D</i> -vine	p=0.49	p=0.93	p=1	p=0.94
	ts=1.43	ts=0.61	ts=0.01	ts=0.61
<i>C</i> -vine	p=0.42	p=0.78	p=1	p=0.84
	ts=1.47	ts=0.68	ts=0.01	ts=0.68
<i>R</i> -vine	p=0.33	p=0.38	p=1	p=0.97
	ts=1.53	ts=4.91	ts=0.01	ts=0.50

The resulting p-values are reported in Table 4.6. None of the structures can be rejected at a 5% significance level, i.e all of them fit the data quite well. As suggested by Gaupp et al. (2017), future developments of improved goodness-of-fit tests are needed to help distinguishing between alternative structuring approaches. We also perform the likelihood ratio tests by Vuong (1989) and

Table 4.7: Vuong test results at level $\alpha = 5\%$

	Statistic	Statistic	Statistic	p-value	p-value	p-value
		Akaike	Schwarz		Akaike	Schwarz
<i>D</i> -vine VS <i>C</i> -vine	1.121	1.637	2.791	0.262	0.201	0.005
<i>R</i> -vine VS <i>C</i> -vine	-1.094	-0.413	1.107	0.273	0.679	0.267
<i>R</i> -vine VS <i>D</i> -vine	-1.641	-1.416	-0.911	0.100	0.156	0.361

Table 4.8: Clark test results

	Statistic	Statistic Akaike	Statistic Schwarz	p-value	p-value Akaike	p-value Schwarz
<i>D</i> -vine VS <i>C</i> -vine	306	311	320	0.181	0.345	0.813
<i>R</i> -vine VS <i>C</i> -vine	293	302	329	0.018	0.098	0.694
<i>R</i> -vine VS <i>D</i> -vine	312	317	338	0.387	0.637	0.270

Clarke (2007) to select the structure that better accounts for the dependence among the assets. The results for each possible couple of vine structures are reported in Table 4.7 and Table 4.8, respectively. These show that, in almost all cases, no decision among the models is possible, i.e the null hypothesis that both models are statistically equivalent cannot be rejected. The only exception is represented by the result of the Clark test where the *R*-vine is preferred to the *C*-vine.

4.4 The optimal composition of long-term natural gas contract

4.4.1 Optimization portfolio problems

In the following, we summarize the optimization portfolio problems that we solve to compute the optimal weights of each underlying asset of a LTC under the minimum portfolio risk. We consider five well-known risk measures that are represented by Variance, Mean Absolute Deviation (MAD), MiniMax, Conditional Value-at-Risk (CVaR), and Conditional Drawdown at Risk (CDaR). We simulate the portfolio returns based on the dependence structures specified in the *C*-vine, *D*-vine and *R*-vine models described above and estimate the risk of the seven-dimensional long-term natural gas contract. We assume to have M assets ($m = 1, \dots, M$) and T time periods ($t = 1, \dots, T$). Recall that, in our analysis, the assets are represented by the seven time series indicated in Section 4.3.1. More precisely, we denote with w_m the weights associated with each asset m of the portfolio; $r_{t,m}$ the return of each asset m in time period t ; μ_m the average return of asset m that is $\mu_m = \frac{1}{T} \sum_{t=1}^T r_{t,m}$ and μ_p the portfolio target return. The mean variance (EV) nonlinear optimization

problem (see Markowitz (1952)) is the following:

$$\min_w \frac{1}{T} \sum_{t=1}^T \left(\sum_{m=1}^M w_m (r_{t,m} - \mu_m) \right)^2 \quad (4.8a)$$

$$s.t. \quad (4.8b)$$

$$\sum_{m=1}^M w_m \mu_m = \mu_p \quad (4.8c)$$

$$\sum_{m=1}^M w_m = 1 \quad (4.8d)$$

$$w_m \geq 0 \quad \forall m = 1, \dots, M. \quad (4.8e)$$

This optimization problem aims at minimizing portfolio variance (4.8a) under the portfolio target return (4.8c), assuming that the sum of the asset weights has to be equal to one (4.8d) and the non-negativity of weights w_m (4.8e).

We then consider the portfolio optimization model that is based on the MAD risk measure (see Konno et al. (1993)):

$$\min_w \frac{1}{T} \sum_{t=1}^T \left| \sum_{m=1}^M (r_{t,m} - \mu_m) w_m \right| \quad (4.9a)$$

$$s.t. \quad (4.9b)$$

$$\sum_{m=1}^M w_m \mu_m = \mu_p \quad (4.9c)$$

$$\sum_{m=1}^M w_m = 1 \quad (4.9d)$$

$$w_m \geq 0 \quad \forall m = 1, \dots, M. \quad (4.9e)$$

Problem (4.9a)-(4.9e) can be transformed in the following linear optimization problem:

$$\min_{w,y} \frac{1}{T} \sum_{t=1}^T y_t \quad (4.10a)$$

$$s.t. \quad (4.10b)$$

$$\left| \sum_{m=1}^M (r_{t,m} - \mu_m) w_m \right| \leq y_t \quad (4.10c)$$

$$\sum_{m=1}^M w_m \mu_m = \mu_p \quad (4.10d)$$

$$\sum_{m=1}^M w_m = 1 \quad (4.10e)$$

$$w_m \geq 0 \quad \forall m = 1, \dots, M. \quad (4.10f)$$

The MiniMax model proposed by Young (1998) aims at maximizing the minimum return L_p , namely minimizing the maximum loss, defined as:

$$L_p = \min_t \left(\sum_{m=1}^M w_m r_{t,m} \right) \quad \forall t = 1, \dots, T.$$

On the basis of this assumption, the model is formulated as follows:

$$\max_{L_p, w} L_p \quad (4.11a)$$

$$s.t. \quad (4.11b)$$

$$\sum_{m=1}^M w_m r_{t,m} - L_p \geq 0 \quad \forall t = 1, \dots, T \quad (4.11c)$$

$$\sum_{m=1}^M w_m \mu_m = \mu_p \quad (4.11d)$$

$$\sum_{m=1}^M w_m = 1 \quad (4.11e)$$

$$w_m \geq 0 \quad \forall m = 1, \dots, M. \quad (4.11f)$$

Following Rockafellar and Uryasev (2000), the portfolio optimization problem with respect to the CVaR measure can be defined as follows:

$$\min_{w, d, v} \frac{1}{(1 - \alpha)T} \sum_{t=1}^T d_t + v \quad (4.12a)$$

$$s.t. \quad (4.12b)$$

$$\sum_{m=1}^M w_m r_{t,m} + v \geq -d_t \quad \forall t = 1, \dots, T \quad (4.12c)$$

$$\sum_{m=1}^M w_m \mu_m = \mu_p \quad (4.12d)$$

$$\sum_{m=1}^M w_m = 1 \quad (4.12e)$$

$$w_m \geq 0 \quad \forall m = 1, \dots, M \quad (4.12f)$$

$$d_t \geq 0 \quad \forall t = 1, \dots, T. \quad (4.12g)$$

where v represents the VaR, $(1 - \alpha)$ is the coverage rate and d_t is the deviation value below the VaR.

Following Chekhlov et al. (2005), the CDaR optimization problem is as follows:

$$\min_{w,u,v,z} \frac{1}{(1 - \alpha)T} \sum_{t=1}^T z_t + v \quad (4.13a)$$

$$s.t. \quad (4.13b)$$

$$\sum_{m=1}^M w_m r_{t,m} + u_t - u_{t-1} \geq 0, \quad u_0 = 0 \quad \forall t = 1, \dots, T \quad (4.13c)$$

$$z_t - u_t + v \geq 0 \quad \forall t = 1, \dots, T \quad (4.13d)$$

$$\sum_{m=1}^M w_m \mu_m = \mu_p \quad (4.13e)$$

$$\sum_{m=1}^M w_m = 1 \dots, M \quad (4.13f)$$

$$w_m \geq 0 \quad \forall m = 1, \dots, M \quad (4.13g)$$

$$z_t \geq 0 \quad \forall t = 1, \dots, T \quad (4.13h)$$

$$u_t \geq 0 \quad \forall t = 1, \dots, T \quad (4.13i)$$

$$(4.13j)$$

where z is an auxiliary vector of variables of the conditional drawdowns, u is the auxiliary vector of variables used to model the cumulative returns and v represents the Drawdown Risk at the quantile $(1 - \alpha)$.

4.4.2 Results

In our analysis we combine pair-vine copula models and portfolio optimization methods to define the optimal allocation of the underlying assets of a LTC. The integration of the PCC into the portfolio optimization allows to capture the complete multivariate dependence risk structure across the considered assets. As mentioned in Section 4.3.2, the PCC exploits the relationship between the pair-copula family and the corresponding Kendall's τ to compute the correlations coefficients among the assets. This methodology does not constraint the returns to be normal, but it captures the asymmetry and nonlinear dependence among the commodities. For each of five risk measures (EV, MAD, MiniMax, CVaR, and CDaR) we minimize the portfolio risk by fixing the same target return μ_p for the three structures. Tables 4.9, 4.10 and 4.11 report the optimal assets allocation for *C*-vine, *D*-vine and *R*-vine structures, respectively. These weights can be interpreted as the proportions to attribute to the different underlyings of LTC, according to the minimum risk optimal portfolio.

Table 4.9: Optimal weights for long-term natural gas portfolio *C*-vine

	EV	MAD	MiniMax	CVaR	CDaR
Brent	0.13	0.14	0.26	0.13	0.38
Gasoil	0.06		0.06		
JetF	0.11	0.17	0.07	0.17	0.62
Naphtha	0.17	0.16	0.12	0.18	
Lsfo	0.31	0.31	0.27	0.30	
Gas NBP	0.12	0.13	0.18	0.13	
Gas HenryHub	0.09	0.09	0.04	0.09	
Min Risk	0.01%	0.03%	0.12%	0.07%	5.32%

The analysis of the optimal asset allocation shows that, in general, there is a convergence in the weight of the same asset within the same risk measure among the three structures. This is in line with the findings of Section 4.3.2, where it is shown that all the three structures are appropriate for modeling the considered series.

The combination of vine copula models with optimization methods leads to optimal portfolios with total risk close to zero.

By focusing on each risk measure, we analyze the optimal asset allocation in the three structures. We first observe that the optimal asset allocation in the three vine copula models is very similar when applying the EV, MAD, and CVaR risk measures (compare Tables 4.9-4.11). In addition, these

Table 4.10: Optimal weights for long-term natural gas portfolio D -vine

	EV	MAD	MiniMax	CVaR	CDaR
Brent	0.16	0.17	0.26	0.17	0.14
Gasoil	0.06				
JetF	0.11	0.17		0.18	
Naphtha	0.16	0.16		0.16	0.19
Lsfo	0.29	0.28	0.58	0.28	0.41
Gas NBP	0.13	0.14	0.16	0.13	0.19
Gas HenryHub	0.09	0.08		0.09	0.07
Min Risk	0.01%	0.03%	0.14%	0.07%	3.67%

Table 4.11: Optimal weights for long-term natural gas portfolio R -vine

	EV	MAD	MiniMax	CVaR	CDaR
Brent	0.16	0.16	0.03	0.17	0.27
Gasoil	0.06				0.59
JetF	0.12	0.19		0.31	0.17
Naphtha	0.14	0.14	0.24	0.14	
Lsfo	0.31	0.29	0.26	0.29	
Gas NBP	0.13	0.13	0.04	0.14	0.01
Gas HenryHub	0.09	0.09	0.11	0.09	0.14
Min Risk	0.01%	0.03%	0.13%	0.07%	4.59%

three risk measures lead to optimal portfolios with equal risk when analyzing the same risk measure. There is also a similar portfolio composition when comparing the aforementioned three risk measures within the same vine structure. Starting from the EV, the resulting portfolio is the one with the minimum risk. All assets are considered with Lsfo constituting approximately 30% of the portfolio, followed by Naphtha, Brent and Gas NBP. Gasoil plays instead a marginal role. This is registered in all the three vine copula models considered. A similar composition results by applying MAD and CVaR risk measures to the three structures. Lsfo is still the asset with the highest weight followed by Jetf, Naphtha and Brent; Gasoil is not included among the optimal underlyings.

The application of MiniMax and CDaR risk measures generates divergences in the components of the optimal portfolio compared to what obtained with the EV, MAD, and CVaR risk measures. A similar behavior for these risk measures is also observed in Bekiros et al. (2015).

We recall that MiniMax considers the maximum loss in the portfolio, while CDaR takes into account a number of draw down events in the historical return distribution. Both measures are sensible to large losses occurring with low probability, which may differ in the simulated distributions of the

three structures. The MiniMax portfolio includes Lsfo, Brent and Gas NBP with different weights in the three structures. In the CDaR portfolio, Brent is the only common underlying among the three structures that show also a significance difference in the total risk. Table 4.12 quantifies the impact of oil-based commodities and spot gases within the optimal portfolios using the information provided by Tables 4.9-4.11. In particular, the values denoted as “Oil-based commodities” are determined by summing up the optimal weights assigned to Brent, Gasoil, JetF, Naphtha, and Lsfo, while the terms “Spot Gases” results from the sum of the optimal weights of Gas NBP and Gas HenryHub.

Table 4.12: Oil and gas composition of the optimal portfolios

Copula	Asset	EV	MAD	MiniMax	CVaR	CDaR
<i>C-vine</i>	Oil-based commodities	0.78	0.78	0.78	0.78	1.00
	Spot Gases	0.22	0.22	0.22	0.22	0.00
<i>D-vine</i>	Oil-based commodities	0.78	0.78	0.84	0.78	0.74
	Spot Gases	0.22	0.22	0.16	0.22	0.26
<i>R-vine</i>	Oil-based commodities	0.78	0.78	0.85	0.78	0.85
	Spot Gases	0.22	0.22	0.15	0.22	0.15

It is worth noting that, when the EV, MAD and CVaR measures are applied, there is a perfect convergence in the composition of the optimal portfolios resulting from the three structures: 78% is constituted by oil-based commodities and the remaining 22% is represented by gas traded on spot markets. On the contrary, the MiniMax and CDaR risk measures lead to optimal portfolios that differ in the three vine copula models, even though the oil-based underlyings still cover the larger share. This confirms the results discussed above.

A synthesis of the portfolio optimization indicates that oil-based commodities, such as Lsfo, Brent, Jetf, and Naphtha, appear to be fundamental picks in our asset allocation, together with the Gas NBP. This is an evidence of the important role still played by the oil-indexation in the LTCs. The influence of oil-based commodities in gas contracts is measured and included in our analysis through the pair-copulas. However, the choice of Gas NBP and Gas HenryHub gases, that together, on average, account for more that 20% of the optimal portfolio, reflects the fact that the re-negotiation policy advocated by mid-streamers in Europe is possible (see Franza (2014)).

4.5 Conclusions

This analysis takes inspiration from the current situation of the European natural gas market where both long-term contracts and hub spot price systems are applied. The fall of the European gas demand combined with the increase of the US shale gas exports and the rise of LNG availability on international markets have led to a reduction of the gas-hub prices in Europe. On the other side, oil-indexed LTCs failed to promptly adjust their positions implying significant losses for European gas mid-streamers that asked for a re-negotiation of their existing contracts and obtained new contracts linked also to hub spot prices. The debate over the necessity of the oil-indexed pricing is still on-going and the main issue is that in the early days of the European gas industry this was the only option to mitigate the risk of launching such a highly capital expensive industry. The supporters of the gas-indexation state that nowadays the gas industry is mature enough and for this reason hub-based pricing reflects the true supply and demand dynamics in natural gas market. This work investigates the risk dependence and the optimal resource allocation of the underlying assets of a LTC through pair-vine copulas and portfolio optimization methods with respect to five risk measures (EV, MAD, MiniMax, CVaR, and CDaR). In order to address the above mentioned debate both spot gas and oil commodities are included as underlyings. The usage of the PCC allows modeling the conditional dependence structure, overcoming the drawbacks of the mean-variance Markowitz optimization, including normally distributed returns and linear correlation among the assets of the same portfolio. The results of our simulations suggest that the weight allocation across portfolios obtained by implementing different vine structures, in almost all cases, converge within the same risk measure. A general finding is that oil commodities still cover the largest share of the optimal portfolios, but spot gas are also included. This suggests that European LTCs will most likely remain indexed to oil-based commodities. In other words, both spot gas and oil-based commodities can be included among the underlyings of long-term gas contract, but the latter will still exercise a major impact. This is a crucial point because increasing the share of spot gas in LTCs would artificially make long-term prices closer to hub price levels. From an economic perspective, this can leave room to a spiral mechanism of downward price adjustment to hub prices that are, in turns, influenced by long-term contract prices.

Appendix

4.A Additional Figures

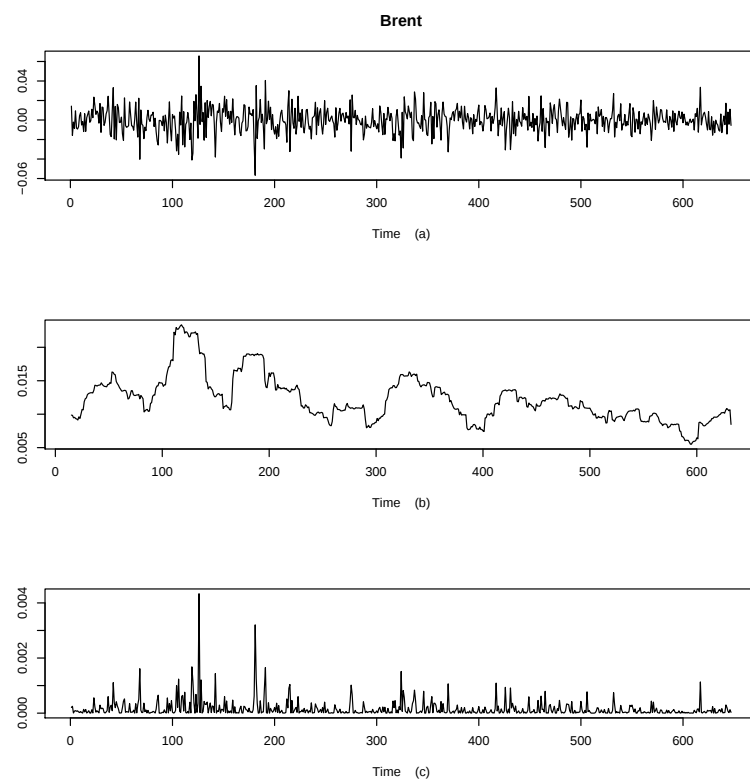


Figure 4.A.1: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Brent series

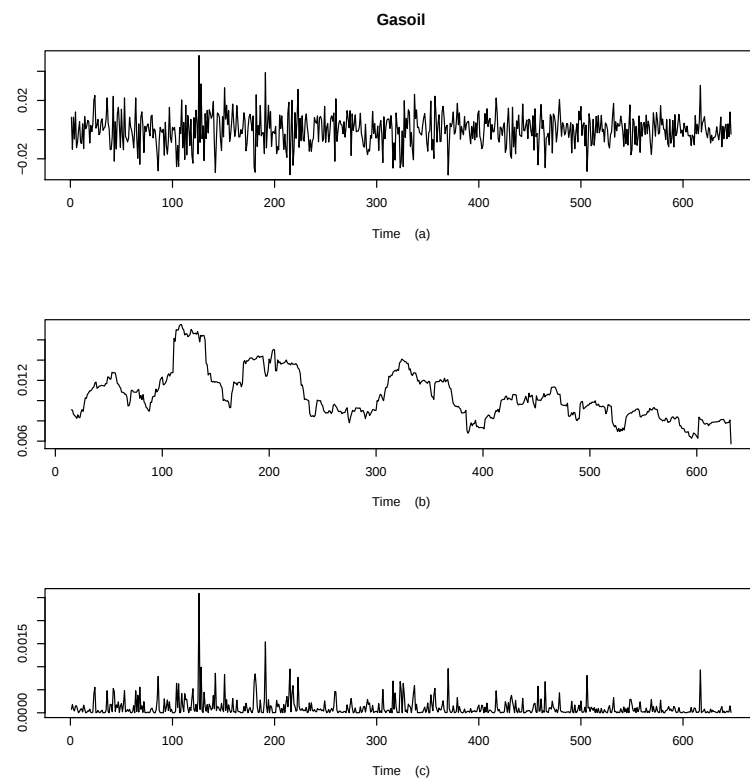


Figure 4.A.2: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Gasoil series.

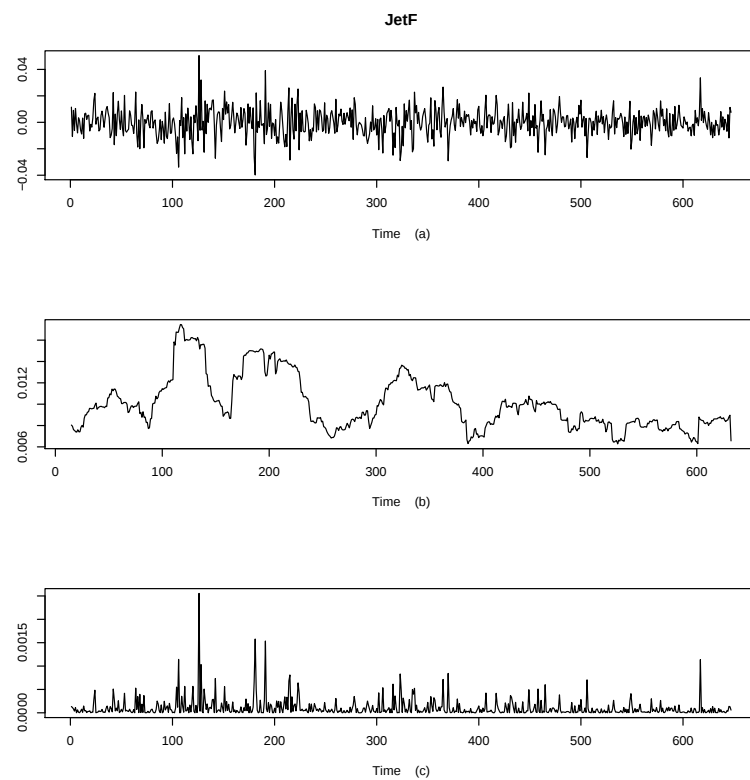


Figure 4.A.3: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the JetF series.

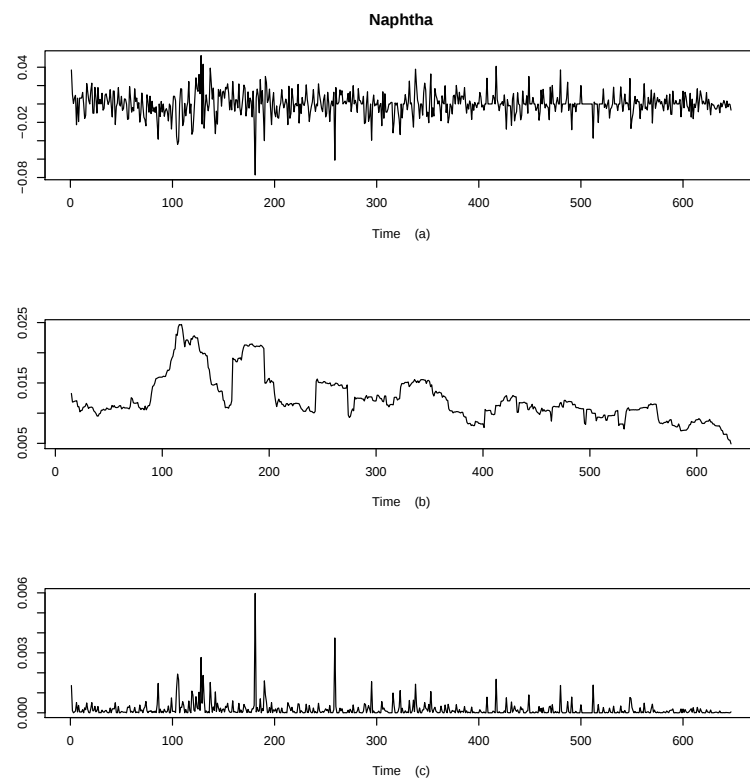


Figure 4.A.4: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Naphtha series.

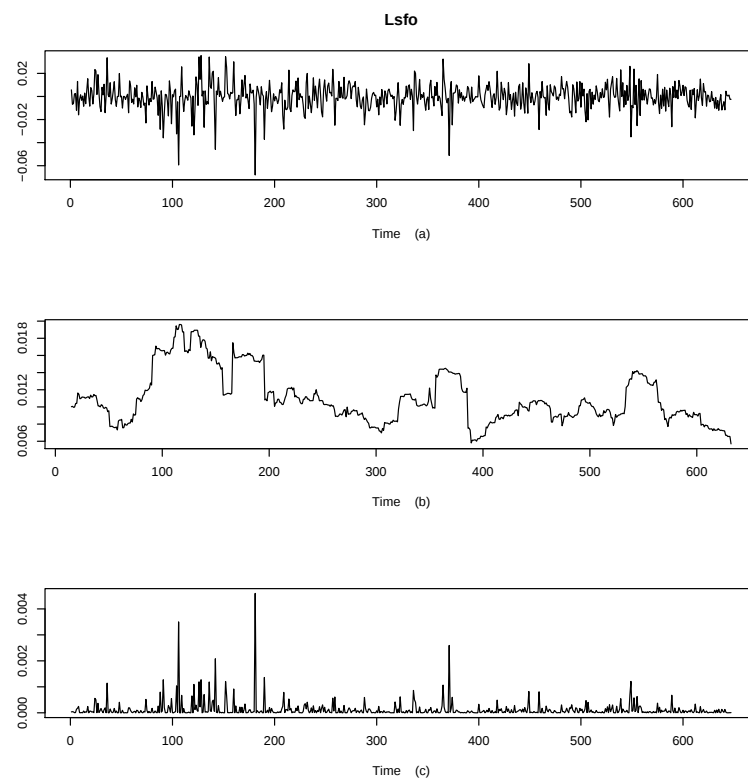


Figure 4.A.5: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Lsfo series.

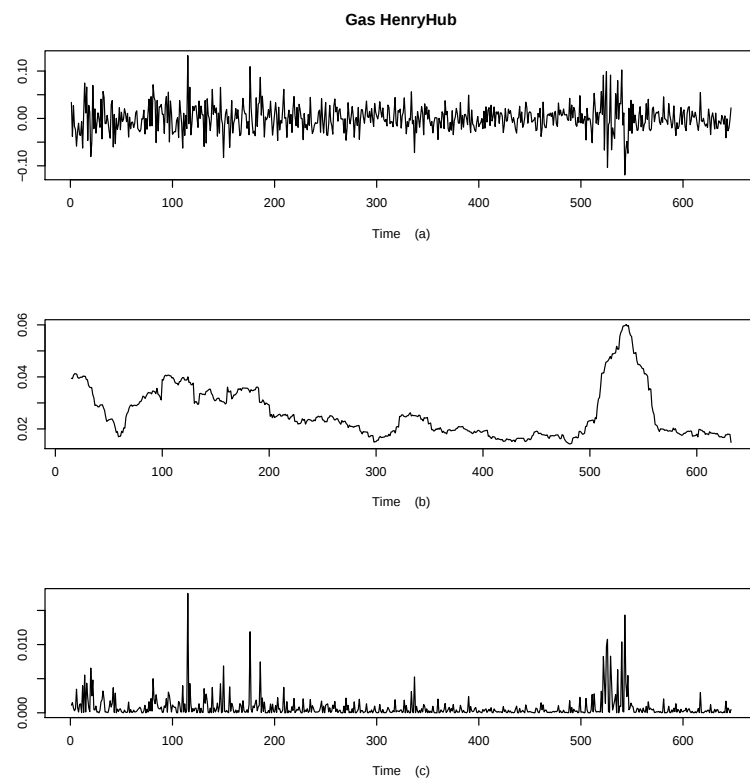


Figure 4.A.6: Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Gas HenryHub series.

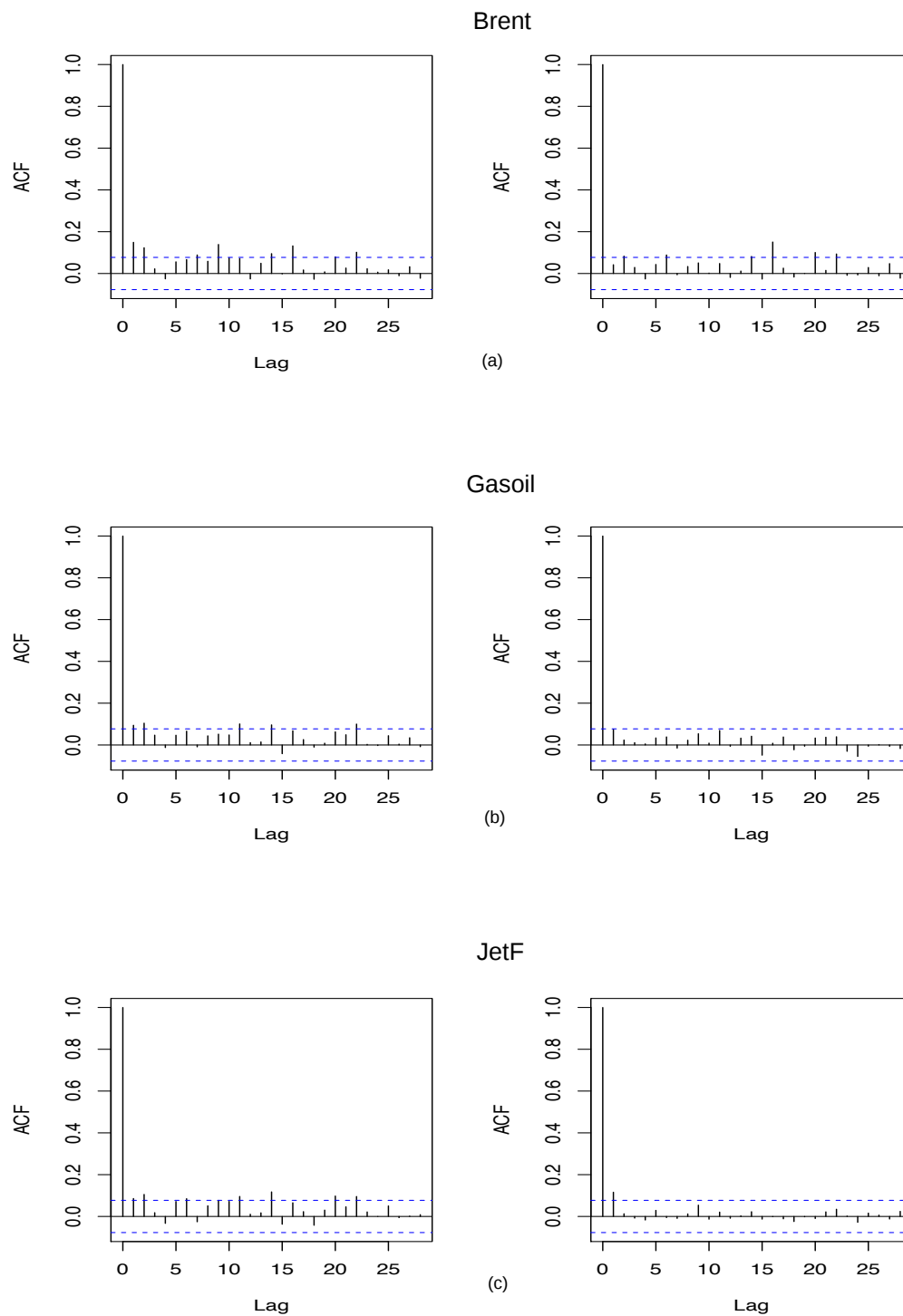


Figure 4.A.7: ACF of the squared mean adjusted log return series and ACF of the squared mean adjusted residuals of Brent (a), Gasoil (b), and JetF (c) log return series.

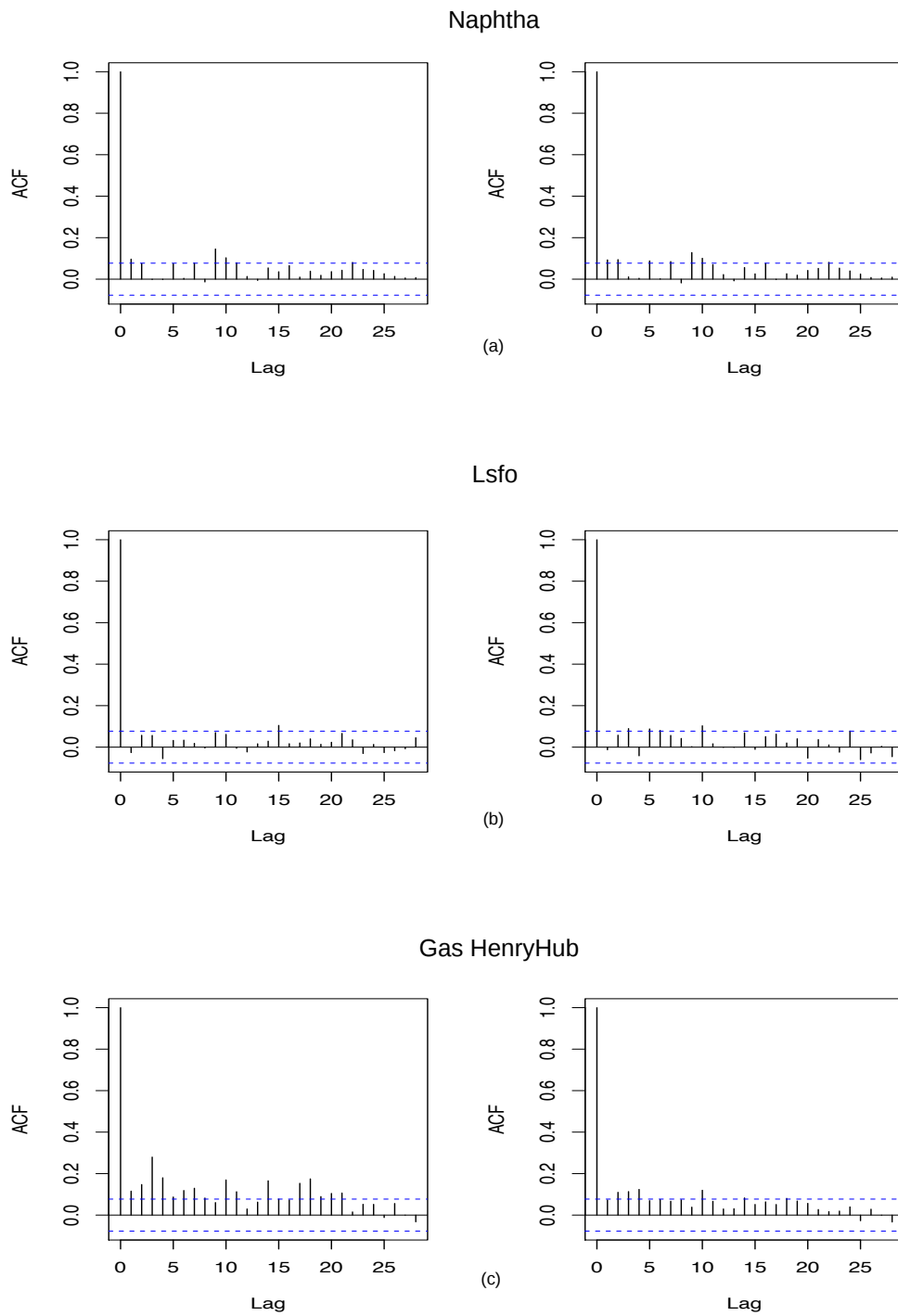


Figure 4.A.8: ACF of the squared mean adjusted log return series and ACF of the squared mean adjusted residuals of Naphtha (a), Lsfo (b), and Gas HenryHub (c) log return series.

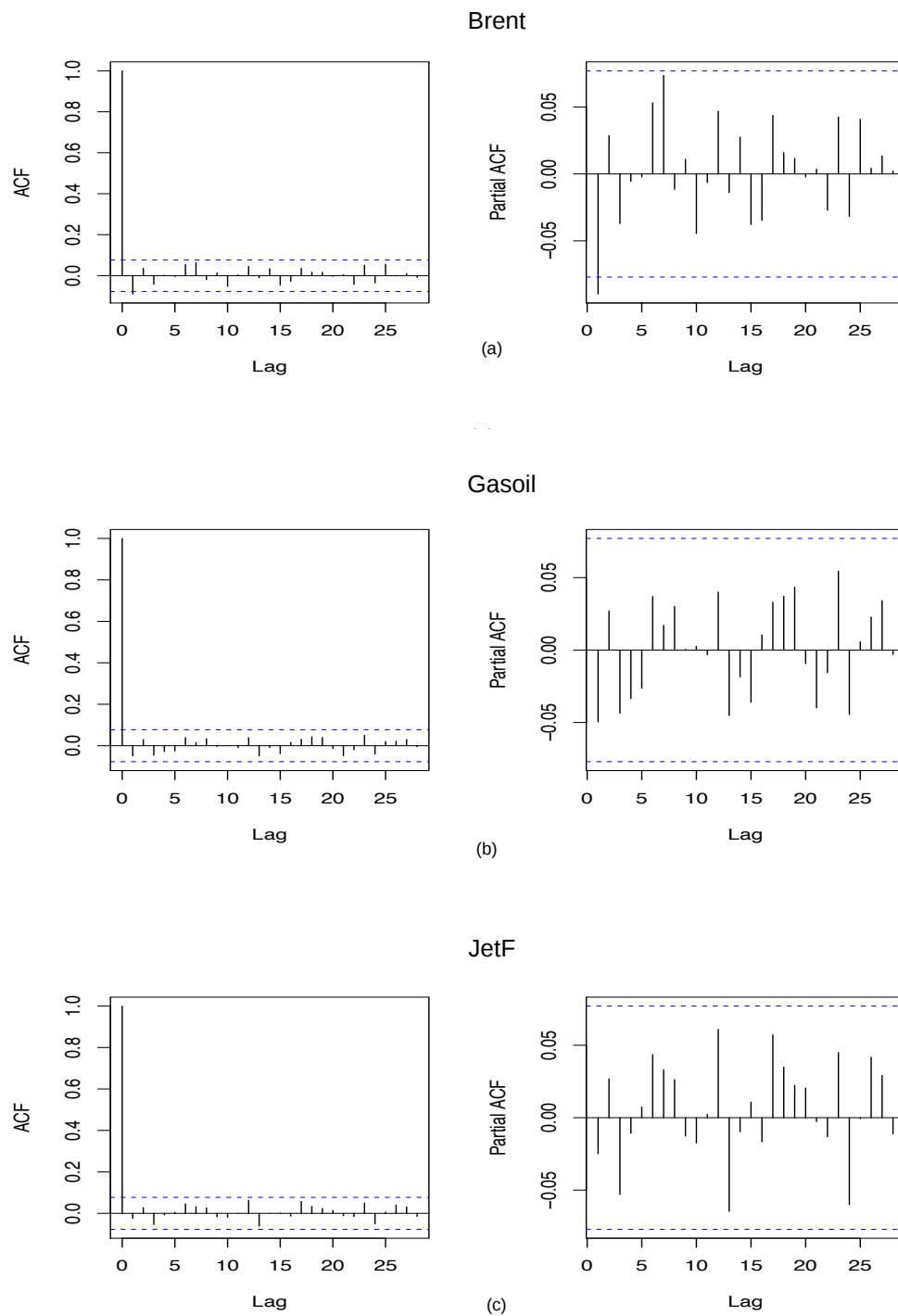


Figure 4.A.9: ACF and PACF of Brent (a), Gasoil (b), and JetF (c) log return series

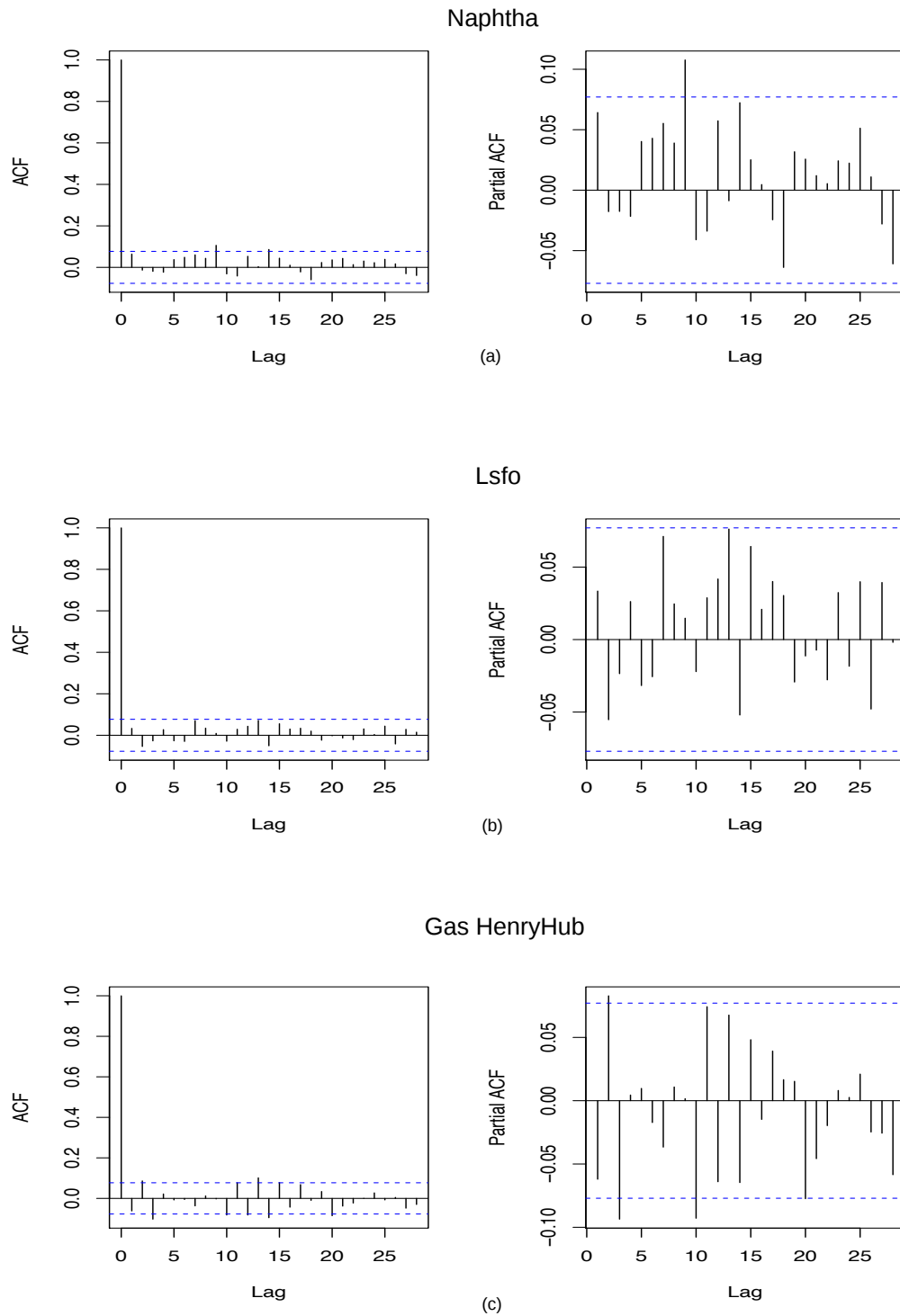


Figure 4.A.10: ACF and PACF of Naphtha (a), Lsfo (b), and Gas HenryHub (c) log return series.

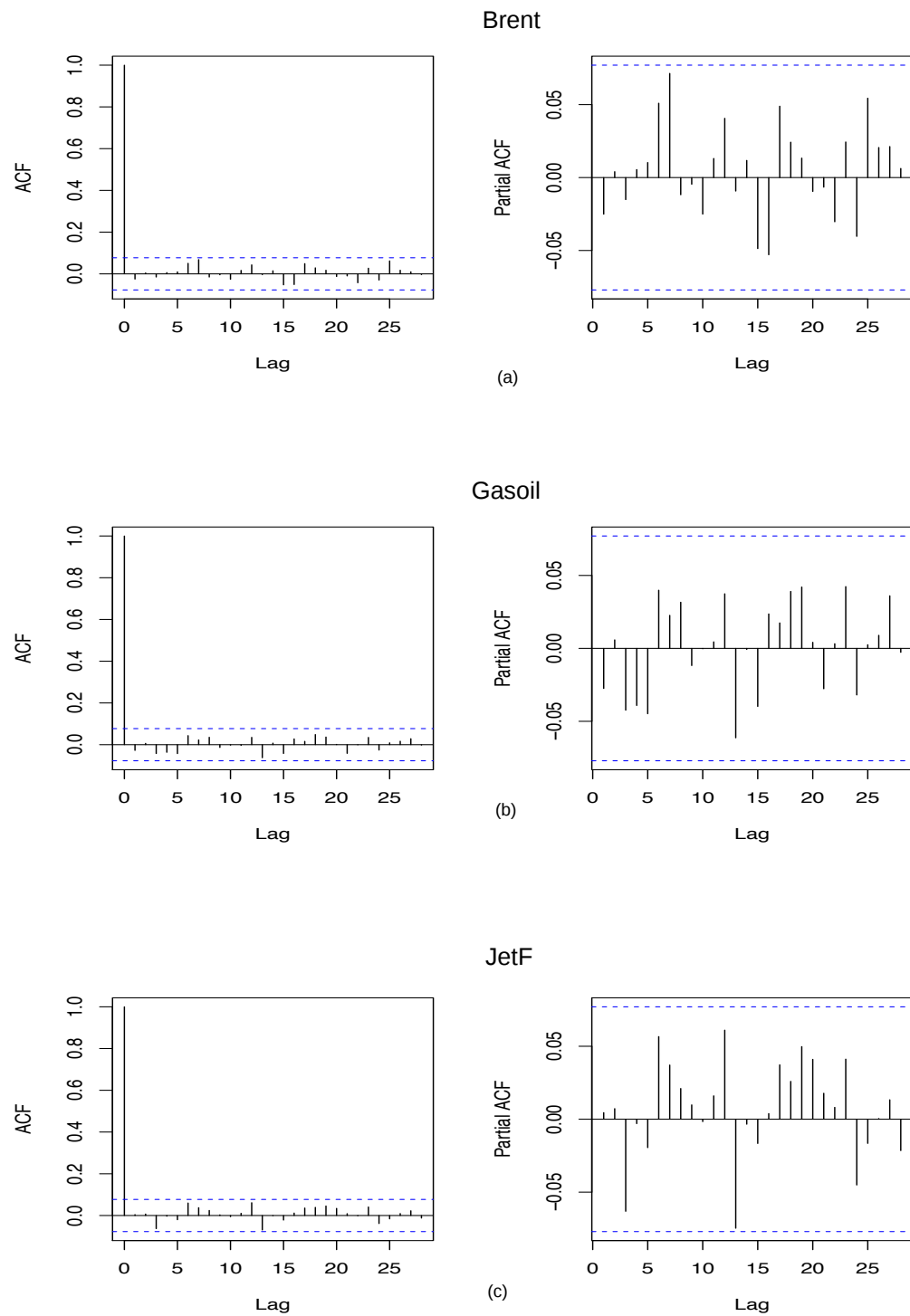


Figure 4.A.11: ACF and PACF of Brent (a), Gasoil (b), and JetF (c) residuals.

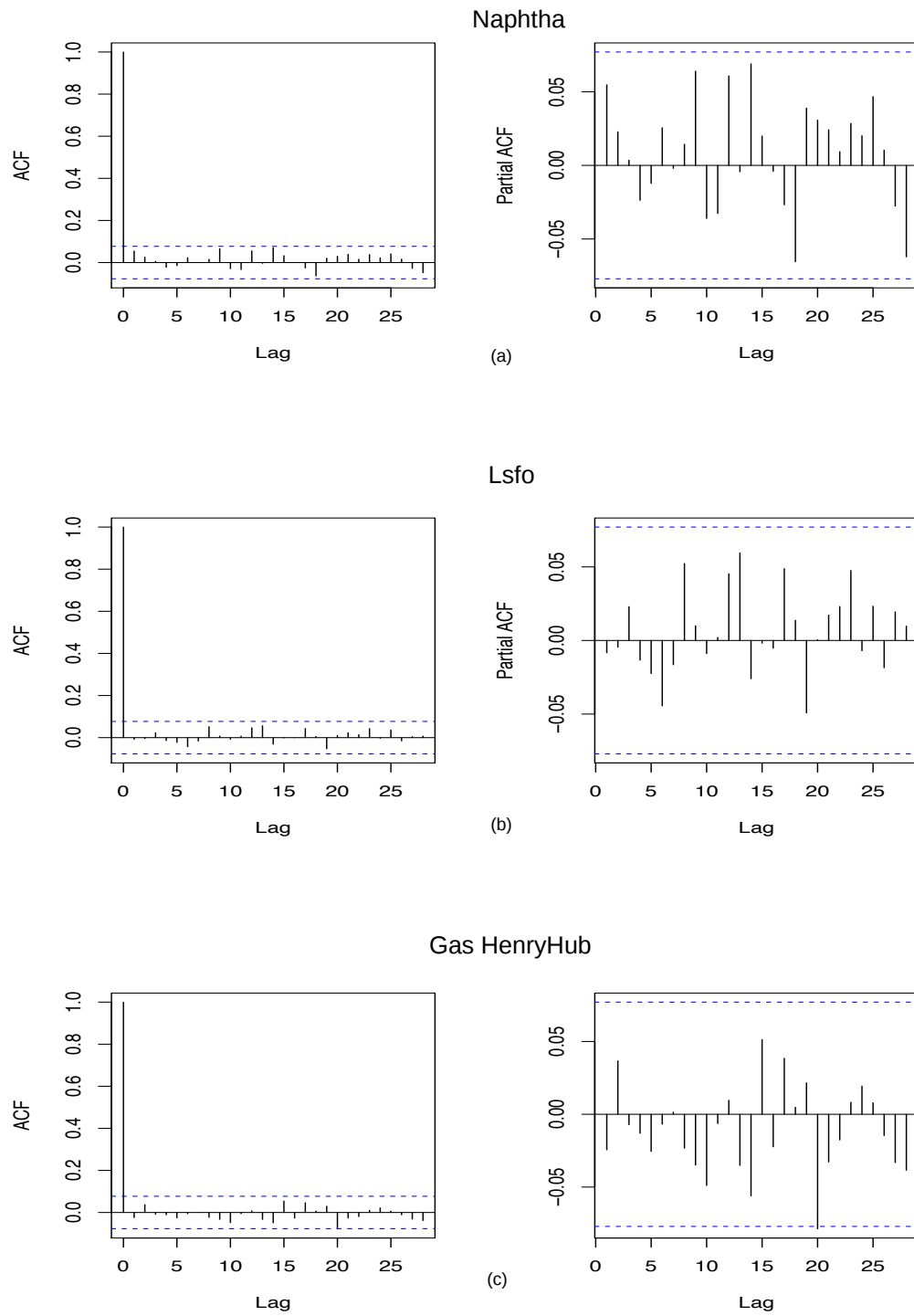


Figure 4.A.12: ACF and PACF of Naphtha (a), Lsfo (b), and Gas HenryHub (c) residuals.

Chapter 5

Evaluating the impacts of the external supply risk in a natural gas supply chain: the case of the Italian market

A large part of the European natural gas imports originates from unstable regions exposed to the risk of supply failure due to economical and political reasons. This has increased the concerns on the security of supply in the European natural gas market.

In this chapter, we analyze the security of external supply of the Italian gas market that mainly relies on natural gas imports to cover its internal demand. To this aim, we develop an optimization problem that describes the equilibrium state of a gas supply chain where producers, mid-streamers, and final consumers exchange natural gas and Liquefied Natural Gas. Both long-term contracts (LTCs) and spot pricing systems are considered. Mid-streamers are assumed to be exposed to the external supply risk, which is estimated with indicators that we develop starting from those already existing in the literature. In addition, we investigate different degrees of mid-streamers' flexibility by comparing a situation where mid-streamers fully satisfy the LTC volume clause ("No FLEX" assumption) to a case where the fulfillment of this volume clause is not compulsory ("FLEX" assumption).

Our analysis shows that, in the "No FLEX" case, mid-streamers do not significantly change their supplying choices even when the external supply risk is considered. Under this assumption, they face significant profit losses that, instead, disappear in the "FLEX" case when mid-streamers are more flexible and can modify their supply mix. However, the "FLEX" strategy limits the gas avail-

ability in the supply chain leading to a curtailment of the social welfare.

The model developed in this chapter have been implemented in GAMS and solved with PATH.

5.1 Introduction

Natural gas covers a significant quota of the energy mix of many of the European countries with a share of 24% in the Total Primary Energy Supply (see Holz et al. (2014)). Many of the European natural gas imports originate from unstable regions and suppliers exposed to the risk of supply failure due to political and economical instability. The political instabilities of Northern Mediterranean area and the disturbances between the European-Russian relationship in the last years have increased the concerns on security of supply in the European natural gas market. Italy is one of the three largest gas consumers in Europe after Germany and the United Kingdom,¹ but it mainly relies on natural gas imports to cover its internal demand since its national production is very low. In particular, the 90.6% of the Italian gas demand in 2015 was satisfied with imports from Russia, Algeria, Libya, the Netherlands, Qatar, and Norway (see Autorità per l'Energia Elettrica il Gas e il Sistema Idrico (2016)).

Considering this framework, in this work we aim at analyzing the security of the external natural gas and Liquefied Natural Gas (LNG) supply (imports) of the Italian gas market. For this reason, we develop an optimization model that describes the equilibrium state of the natural gas supply chain where natural gas producers (suppliers), mid-streamers, and consumers can sell and buy both natural gas and LNG through long-term contracts (LTCs) or/and on spot market. Mid-streamers take on the role of intermediates in this supply chain network: on one side they exchange natural gas and LNG with supplying countries; on the other side they sell gas to final consumers. Since they are in charge of selecting the origin of natural gas and LNG imports, we assume that mid-streamers are the agent group exposed to the external supply risk associated with the imports from foreign countries. In other words, mid-streamers define the amount of gas and LNG to be imported not only on the basis of the relative production and transportation costs, but also taking into account the external supply risk related to the countries from which the gas originates. This risk is evaluated through indicators that we develop starting from those already existing in the literature (see Section 5.3.2 for a literature review on these indicators). These risk indicators are then inserted in the optimization problem as a weight assigned to the Italian natural gas and LNG

¹See Eurostat at http://ec.europa.eu/eurostat/statistics-explained/index.php/Natural_gas_consumption_statistics

imports operated by mid-streamers.

Several works in the literature studies different aspects of the natural gas market. For instance, Egging and Gabriel (2006) develop a mixed complementarity model for the European natural gas market where producers can behave either in a perfectly competitive or in a Cournot strategic way, while the other players in storage and transmission services operate in perfect competition only. This model investigates the impacts of producers' market power and the importance of pipeline and storage capacity. Egging et al. (2008) propose a detailed complementarity model for defining the equilibrium of the European natural gas market under a set of scenarios, including the disruption of Ukrainian pipelines, the disruption of Algerian supplies, and the increased transportation costs. Market players are represented by producers, pipeline and storage operators, marketers, liquefiers and regasifiers, LNG tankers, and final consumers. Egging et al. (2010) present an extensive model of the global natural gas market that allows for the description of the flows and endogenous investment decisions in infrastructures and the evaluation of the market power in the pipeline and LNG markets. Holz et al. (2016) analyze the infrastructure needs of the European natural gas market to adequate it to the decarbonization targets of the European energy system. Their results show that the current import infrastructures and intra-European transit capacity are sufficient to accommodate future import needs under the scenarios of increasing and decreasing consumption. The authors conclude that the supply security would benefit from relaxing the political and technical constraints on investments. Egging and Holz (2016) develop a stochastic global gas market model to study the infrastructure investments and the trade in an imperfect market structure taking into account the possible disruption of the Russian gas transit via Ukraine from 2020; the variation of the electricity intensity generation in the OECD countries after 2025; the availability of the US shale gas after 2030.

All these papers analyze the possible supply disruption with a set of scenarios, operating a sensitivity analysis. In our work, we depart from this approach and we study the risk of external supply by directly integrating new risk indicators in our model. Our goal is to detect whether this risk affects the import choices of mid-streamers that can change not only their supplying country mix but also the type of gas purchased (i.e. they can favor gas to LNG or viceversa) and the payment method (i.e. LTCs vs. spot). To the best of our knowledge, it is the first time that such an approach is considered and this represents one of the contributions of our work.

A second contribution of this work is represented by the investigation of two different degrees of mid-streamers' flexibility. In particular, we compare a situation where the mid-streamers fully accomplish the LTC volume requirements with a case where the mid-streamers behave in a more

flexible way and they are not obliged to purchase all the quantity of gas and LNG contracted with LTCs. We assume that this flexibility regards both the gas and the LNG bought through LTCs, with and without external supply risk. This analysis aims at describing the current configuration of the European gas market where the co-existence of LTCs and hub-pricing systems implies that the natural gas is traded at two different prices on the same market. Depending on the conditions, spot price can be higher or lower than long-term contractual gas, implying possibly difficult situations for companies loaded with high price TOP gas against the low spot prices. This is what happened in the last years, where an excess of gas availability on the market has been reflected into prices at the European hubs lower than those fixed in the LTCs. As a reaction, mid-streamers have asked for an increase of the LTCs flexibility (see Sections 5.2 and 3.5 for more details).

For our analysis, we consider a single optimization problem that is then transformed in complementarity form through the corresponding Karush-Kuhn-Tucker conditions (see Facchinei and Pang (2007); Gabriel et al. (2013); Nagurney (1999)). Complementarity-based models facilitate the formulation of equilibrium problems that describe the interactions of several agents whose choices are subject to technical and economic constraints, as it happens in our model.

The remainder of the chapter is organized as follows. Section 5.2 illustrates the natural gas supply chain that we consider in our analysis and we provide some insights related to mid-streamers' behaviour in the European gas market. Section 5.3 describes the new external supply risk indicators that we construct and the optimization models that we develop. The case study is presented in Section 5.4, while Section 5.5 presents the results of our analysis. Section 5.6 concludes with the final remarks.

5.2 The natural gas market

In this work, we aim at analyzing the interactions of the main players of a natural gas supply chain taking into account the external supply risk. The considered players are N supplying countries (producers), M mid-streamers, and S consumers. The latest two agents' groups are both located in the destination country that, in our case, is represented Italy (see 5.1).

We assume that gas can be sold/purchased either through long-term contracts (LTCs) or on spot markets (hub pricing). This assumption aims at representing the current situation of the European gas market where LTCs and the hub pricing systems co-exist, even though it is still dominated by long-term contracts. The LTCs have been traditionally concluded over long periods (typically

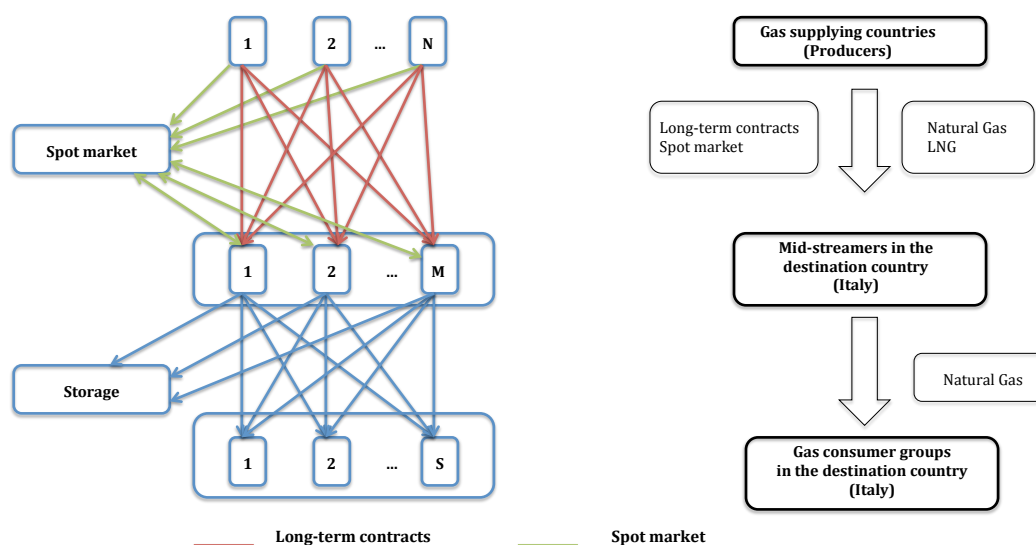


Figure 5.1: Natural gas supply chain

20 years or more) and are characterized by quantity and price clauses that have been historically introduced to allow for risk sharing between gas producers and mid-streamers that respectively face price and volume risks (Abada et al. (2017)). The Take or Pay (TOP) quantity clause obligates the buyer to take a certain quantity of natural gas or to pay for it. The Price Indexation clause relates the price at which gas is bought to some index on the market that has been traditionally represented by the price of crude and oil-products. The hub pricing approach developed in the nineties in the US and UK and is now developing in Europe. In this system, natural gas is traded, every day, on a spot market that determines prices and volume on the short term. International natural gas market is organized in different ways depending on the considered areas. North America is essentially organized on the basis of Henry Hub spot market; Asia is mainly supplied through long-term contracts; Europe is still dominated by long-term contracts, even though spot markets are growing and are expected to develop further. The main exception in Europe is represented by the UK where gas is largely traded at the National Balancing Point (NBP) spot market. In continental Europe, Zeebrugge (ZEE) and the Title Transfer Facility (TTF), respectively located in Belgium and in the Netherlands, are the two dominant spot market places and many others are emerging such as the Punto di Scambio Virtuale (PSV) in Italy (see Melling (2010)).

The co-existence of two pricing systems in Europe implies that the natural gas is traded at two different prices on the same market, causing possibly difficult situations for companies (mid-streamers) that are charged with high LTC prices against the low spot prices. This is what happened in the last

years, where the combined effects of the increase of the US shale gas exports, the reduction of European gas demand due to the economic crisis, and the increased availability of uncommitted LNG from Qatar led to an excess of gas availability on the market that was reflected into low gas prices at the European hubs. On the other side, oil-indexed long-term gas contracts failed to promptly adjust their positions implying significant losses for European gas mid-streamers that were stuck with their LTCs and could only dump the excess of gas on the spot market. As a consequence of this short-term but dramatic issue, European mid-streamers have asked for a re-negotiation of their long-term gas contracts to make them more flexible and closer to spot gas prices. These re-negotiations have resulted into a decline of oil-indexation and hub-linked pricing has rapidly become the basis for an increasing number of transactions in the European gas market. In addition, with regard to newly signed contracts for pipeline sales to Europe, there is a clear trend towards shorter commitments (see Franza (2014)).

In addition to the external supply risk, this work aims at analyzing this structural problem of the gas industry by focusing on the short term and considering different degrees of flexibility of mid-streamers' behaviour. In particular, we compare the following two cases:

- **No flexibility:** In this case, we assume that the mid-streamer has to comply with all the LTCs that it already has. This corresponds to the situation where the mid-streamer has to buy at least the amount of natural gas and LNG already contracted. Under this assumption, they are also allowed to conclude new LTCs.
- **Flexibility:** This case aims at representing the current situation of the European gas market where LTCs and the hub pricing systems co-exist. In particular, we assume that the mid-streamer has the possibility to decide whether or not buying gas or LNG via LTCs. In other words, the mid-streamers is allowed to not respect the TOP clause for a short period (as the time framework that we consider).

As depicted in Fig. 5.1, in our model, we assume that supply countries produce gas and LNG that they directly sell to mid-streamer with LTCs or on the spot market. On the other side, mid-streamers can decide to buy gas and/or LNG from supplying countries through LTCs. They can also operate on the spot market by either buying and selling gas. On the other side, we assume that mid-streamers can only buy LNG on spot because according to International Group of Liquefied Natural Gas Importers (2018), Italy does not re-export the imported LNG. Mid-streamers are also in charge of managing the storage site in the destination country. Finally, we consider three groups of consumers represented by industry, power generation, and residential/commercial.

5.3 Modeling the gas supply chain with external supply risk

In this section, we develop the optimization model used to describe the gas supply chain with producing countries, mid-streamers operating, and consumers located in the destination country. We consider a time span of one year, sub-divided into two time segments corresponding to a low-demand and high-demand periods. We start from the notation used in the mathematical formulation, we then describe the external supply risk indicators, and finally we present the optimization model.

5.3.1 Notation

Indices

- N : number of countries producing and supplying natural gas (producers/supplying countries), $n = 1, \dots, N$;
- M : number of mid-streamers located in the destination country, $m = 1, \dots, M$. In our model, we assume one destination country;
- F : number of natural gas entry points located in the destination country, $f = 1, \dots, F$;
- S : number of consumption sectors in the destination country (industry, power generation, and residential/commercial) $s = 1, \dots, S$;
- T : number of time periods, $t = 1, \dots, T$. More precisely, we consider $T = 2$, where $t = 1$ is low-demand period and $t = 2$ is high-demand period.

Parameters:

- θ_t : duration in days of time periods t ;
- $\bar{X}_n$: gas production capacity of producer n ;
- $\bar{GP}_n$: capacity of gas pipeline connecting producing country n with the destination market where mid-streamers are located;
- $\bar{L}_n$: liquefaction capacity of producer n ;
- $\bar{R}_m$: regasification capacity of mid-streamer m ;
- $\bar{I}_m$: injection limit of storage site managed by mid-streamer m ;

- $\bar{W}_m$: withdrawal limit of storage site managed by mid-streamer m ;
- WG_m : working gas volume available at storage site managed by mid-streamer m ;
- α_n : rate of liquefaction loss faced by producer n ;
- β_m : rate of regasification loss faced by mid-streamer m ;
- stc_{nm}^{LNG} : transportation costs of LNG through ship from the producing country n to the destination market where mid-streamers m are located;
- ptc_{nm}^G : transportation costs of gas through pipelines from the producing country n to the destination market where mid-streamers m are located;
- ptc_n^{SpotG} : pipeline transportation costs of gas sold on spot faced by the producing country n ;
- dc_{ms} : gas distribution costs through pipelines faced by mid-streamer m to supply consumer s ;
- τ_{nm} : (minimum) annual amount of gas that producer n supplies to mid-streamer m through long-term contracts;
- ξ_{nm} : (minimum) annual amount of LNG that producer n supplies to mid-streamer m through long-term contracts;
- Υ_{ft} : limit of the pipeline entry point f in the destination country in time period t ;
- Γ_{fn} : limit of the pipeline entry point f that receives gas from producing country n in time period t .

Variables

- X_{nt}^G : total amount of natural gas produced by supplying country n in one day of time period t (mcm/d).
- x_{nmt}^G : amount of natural gas supplied by producer n through long-term contracts (LTCs) to mid-streamer m in one day of time period t (mcm/d).
- X_{nt}^{LNG} : total amount of natural gas transformed in LNG by supplying country n in one day of time period t (mcm/d).
- x_{nmt}^{LNG} : amount of LNG supplied by producer n through long-term contracts (LTCs) to mid-streamer m in one day of time period t (mcm/d).

- x_{nt}^{SpotG} : amount of natural gas sold by producer n on spot market in one day of time period t (mcm/d).
- $x_{nmt}^{SpotLNG}$: amount of LNG sold by producer n to mid-streamer m on spot basis in one day of time period t (mcm/d).
- Y_{mt}^{LNG} : total amount of natural gas re-gasified by mid-streamer m in one day of time period t (mcm/d).
- y_{nmt}^{LNG} : amount of LNG purchased by mid-streamer m through LTC from producer n in one day of time period t (mcm/d).
- y_{mt}^{SpotG} : amount of natural gas purchased by mid-streamer m on spot market in one day of time period t (mcm/d).
- $y_{nmt}^{SpotLNG}$: amount of LNG purchased by mid-streamer m from producer n on spot basis in one day of time period t (mcm/d).
- q_{mt}^{SpotG} : amount of natural gas sold by mid-streamer m on spot market in one day of time period t (mcm/d).
- z_{mst} : nonnegative amount of natural gas that mid-streamer m sells to consumer sector s in one day of time period t in Bcm.
- i_{mt} : nonnegative amount of natural gas injected by mid-streamer m in the storage site in one day of time period $t = 1$ in Bcm.
- w_{mt} : nonnegative amount of natural gas withdrawn by mid-streamer m from the storage site in one day of time period $t = 2$ in Bcm.
- d_{st} : nonnegative amount of natural gas demanded by consumer s in one day of time period t in Bcm.
- $P_{st}(d_{st})$: Inverse demand function of consumer s in one day of time period t in Bcm. This function can be stated as $P_{st}(d_{st}) = a_{st} - b_{st} \cdot d_{st}$ where a_{st} is the intercept of consumers' (affine) demand functions in time period t (€/Bcm) and b_{st} is the slope of consumers' (affine) demand functions in time period t (€/Bcm²).

All these variables are assumed to be nonnegative. As already indicated, we consider a time span of one year subdivided into low-demand and high-demand periods with a duration in days θ_t , respectively. For each of these two periods, we consider a representative day and, therefore, variables

and parameters have to be weighted by the duration θ_t in order to get annual values. Finally, we do not list here the dual variables associated with the constraints appearing in our model formulation. These are directly indicated next to the constraints to which they refer.

5.3.2 External supply risk indicators

Energy security is defined as the availability of a regular supply of energy at an affordable price (see International Energy Agency & Organisation for Economic Co-operation and Development (2001)). Security of gas supply in energy systems has always been an important issue due to the high dependence on energy. This is particularly true for Europe where about one quarter of all the energy used is natural gas, and many European countries import nearly all their supplies, as it happens for Italy. Supply disruptions caused by infrastructure failure or political disputes are real phenomena. As an example, we recall the the severe shortfall of gas in Western Europe due to Russia's decision to suspend gas deliveries to Ukraine in January 2009. A way to deal with energy security is a process of managing the associated risk. For this reason, in our analysis, we concentrate on short-term indices to assess the risk associated with external energy supply and possible insecurity of supply. The Herfindahl-Hirschman Index (HHI), the Shannon-Wiener Index (SWI), and the variations of it, such as the Shannon-Wiener-Neumann indeces (SWNIs) are amongst the most commonly used aggregate indicators for energy security applied to evaluate the diversification of the market (see Neumann (2004)).

The HHI is adopted as a measure of market concentration, namely the total number of companies operating. In a similar way, the diversity of energy suppliers is given by the sum of the squares of each supplier market share q_n :

$$HHI = \sum_{n=1}^N q_n^2 \quad (5.1)$$

where q_n represents the share of imports from a particular country n into the country considered. Thus, the higher the value of the index, the more concentrated the market is; the maximum value of the index is achieved when there is only one supplier. Consequently, the HHI is only used to investigate markets with a lack of diversity of suppliers, or to give greater weight to the larger suppliers (see e.g. Blyth and Lefevre (2004); Cohen et al. (2011); Grubb et al. (2006); Gupta (2008); Kruyt et al. (2009); Frondel and Schmidt (2008)).

The Shannon-Wiener concentration Index is an alternative approach to measure the diversity of

energy suppliers. This index is computed as follows:

$$SWI = - \sum_{n=1}^N q_n \ln q_n \quad (5.2)$$

where q_n is the supplier's market share as in the HHI. The higher the value of the index, the more diverse the market is. Moreover, differently from the HHI, it puts more weights on the impact of smaller suppliers (see Kruyt et al. (2009), footnote 13, for more details on the mathematical properties of the SWI).

However, the security of external energy supply may be affected by the political situation in the exporting country. Starting from the SWI, Neumann (2004) takes this into account by using a measure of the political stability of the supplier n that we denote with r_n . In particular, Neumann (2004) proposes two additional indicators $SWNI_1$ and $SWNI_2$ that are defined as follows:

$$SWNI_1 = - \sum_{n=1}^N (r_n q_n \ln q_n), \quad (5.3)$$

and:

$$SWNI_2 = - \sum_{n=1}^N (r_n q_n \ln q_n)(1 + g_n), \quad (5.4)$$

where r_n identifies the political risk rating associated with the supplying country n and g_n in the $SWNI_2$ represents the indigenous production of the resource in the supplying country n .

On the other side, Quemada et al. (2012) propose the Geopolitical Energy Security (GES) indicator that is computed as:

$$GES = \sum_{f=1}^F \left[\left(\sum_{n=1}^N r_n q_{nf}^2 \right) \cdot e^{\frac{1}{P_f}} \right] \cdot \frac{C_f}{TPES} \quad (5.5)$$

where the index f represents the fuel type, and the ratio $\frac{C_f}{TPES}$ denotes the share of total consumption of fuel f in the Total Primary Energy Supply (TPES). The GES indicator combines a country exposure measure of market concentration q_{nf}^2 , similarly to the HHI, with the market liquidity, involving as a key role the exponential function e^{1/P_f} where P_f represents the ratio between the offer of the resource f and its consumption. The GES indicator is an expansion of the ESI_{price} index that does not consider the market liquidity, i.e. the exponential function (see International Energy Agency (2007)).

Le Coq and Paltseva (2009) develops the Risky External Energy Supply (REES) for each destination country a that is defined as follows:

$$REES_a^f = \left[\sum_{n=1}^N \left(\frac{NPI_{na}^f}{NPI_a^f} \right)^2 F_{na}^f r_n d_{na} \right] \cdot NID_a^f \cdot SF_a^f \quad (5.6)$$

where NPI_{na}^f is the net positive imports of fuel f from supplying country n to the destination country a , NPI_a^f is the sum of the net positive imports over all suppliers of country a , F_{na}^f is the fungibility of fuels f imported from n to a , r_n is the political risk index of the supplier country, d_{na} is a measure of a distance between countries n and a , NID_a^f is the net import dependency of country a for fuel f and SF_a^f is a share of fuel f in country a (see Le Coq and Paltseva (2009)).² The REES is the most comprehensive among the indicators here illustrated, since it accounts for the diversification of the energy sources f , the political risk r_n , the distance between supplying and destination country d_{na} , and the fungibility F_{na}^f of imported fuels, in addition to the diversification of suppliers and the energy dependence. Note that the distance d_{na} represents a proxy for the risk involved in the transportation phase. On the other side, the market liquidity is a feature of the GES indicator.

Our analysis aims at measuring the effects of the external supply risk on the natural gas and LNG exchanges between producing and supplying countries n and mid-streamers m all operating in same destination country. In this framework, we assume that mid-streamer selects the supply countries from which importing natural gas and LNG on the basis of external supply risk of these producing countries. Our idea is to consider the external supply risk as an additional cost that the mid-streamers face when buy both natural gas LNG from the different supplying countries. This cost is indeed proportional to the amount of gas bought and therefore, in our model formulation, we weight each exchange of natural gas and LNG between the producers and the mid-streamers by the associated external supply risk that varies according to the considered supplying country. We are therefore interested in evaluating a risk per each type of natural gas imported and per each supplying country. The indicators presented above aggregated all this information that we, instead, need in a disaggregated form.

For this reason, we consider the HHI, the SWNI₂, the GES, and the REES presented above and we modify them in order to make them suitable for our scope. In this way, we develop new indicators

²Note that all the other indicators presented above are computed per each destination country even though not explicitly indicated.

that are all denoted as “ Π ”. For their construction, we account for two fuels, natural gas (G) and LNG ($f = G, LNG$ considering the GES and the REES indicators) and one destination country (this means that, using the REES notation, $a = 1$). This latter assumption implies that mid-streamers face an identical external supply risk when importing the same type of gas from the same country.

The new indicators that we introduce are:

1.

$$\Pi_{nm}^{HHI,G} = (q_{nm}^G)^2 \quad \text{and} \quad \Pi_{nm}^{HHI,LNG} = (q_{nm}^{LNG})^2, \quad (5.7)$$

These are a transformation of the HHI where q_{nm}^G and q_{nm}^{LNG} represents the share of imports of mid-streamer m from a particular country n respectively for natural gas and LNG. For the reasons explained above, we do not consider the sum over all supplying countries.

2.

$$\Pi_{nm}^{SWNI_2,G} = -(r_n q_{nm}^G \ln q_{nm}^G)(1 + g_n), \quad (5.8)$$

$$\Pi_{nm}^{SWNI_2,LNG} = -(r_n q_{nm}^{LNG} \ln q_{nm}^{LNG})(1 + g_n), \quad (5.9)$$

These are similar to the $SWNI_2$, but, as in the $\Pi_{nm}^{HHI,G}$ and $\Pi_{nm}^{HHI,LNG}$, are computed taking into account the share of natural gas and LNG that country n supplies to mid-streamer m , without considering the sum over all supplying countries. The other terms r_n and g_n are as in the $SWNI_2$ indicator.

3.

$$\Pi_{nm}^{GES,G} = \left[\left(r_n (q_{nm}^G)^2 \right) \cdot e^{\frac{1}{P_G}} \right] \cdot \frac{C_G}{TPES}, \quad (5.10)$$

$$\Pi_{nm}^{GES,LNG} = \left[\left(r_n (q_{nm}^{LNG})^2 \right) \cdot e^{\frac{1}{P_{LNG}}} \right] \cdot \frac{C_{LNG}}{TPES}, \quad (5.11)$$

These result from the modification of the GES indicator, where we consider the share of natural gas and LNG imported by the mid-streamer m from country n , without operating the

sum overall supplying countries. In addition, we maintain separate the two fuels supplied. The other terms are as in the GES indicator.

4.

$$\Pi_{nm}^{REES,G} = \pi_{nm}^G \cdot r_n \cdot d_{nm} \cdot F_{nm}^G, \quad (5.12)$$

$$\Pi_{nm}^{REES,LNG} = \pi_{nm}^{LNG} \cdot r_n \cdot d_{nm} \cdot F_{nm}^{LNG} \quad (5.13)$$

where, as in the REES indicator, r_n is the measure of political risk, d_{nm} is a factor that accounts for the distance between the capitals of the producing country and the location of mid-streamers, F_{nm}^G and F_{nm}^{LNG} are the fungibility respectively of natural gas and LNG, and π_{nm}^G and π_{nm}^{LNG} are the shares of natural gas and LNG that mid-streamer m in the destination country imports from supplying country n . These shares are computed as indicated below:

$$\pi_{nm}^G = \left(\frac{\tilde{q}_{nm}^G}{\sum_{n=1}^N \tilde{q}_{nm}^G} \right)^2$$

$$\pi_{nm}^{LNG} = \left(\frac{\tilde{q}_{nm}^{LNG}}{\sum_{n=1}^N \tilde{q}_{nm}^{LNG}} \right)^2$$

where $\tilde{q}_{nm}^G$ and $\tilde{q}_{nm}^{LNG}$ are parameters defining the *net* gas and LNG imports of mid-streamer m from the supplying country n .

Table 5.1 summarizes the indicators already existing in the literature, which we have present above, and the new indicators that we introduce starting from them.

Table 5.1: Summary of the new indicators proposed

Indicators existing in the literature	New indicators proposed	
HHI	$\Pi_{nm}^{HHI,G}$	$\Pi_{nm}^{HHI,LNG}$
SWNI ₂	$\Pi_{nm}^{SWNI_2,G}$	$\Pi_{nm}^{SWNI_2,LNG}$
GES	$\Pi_{nm}^{GES,G}$	$\Pi_{nm}^{GES,LNG}$
REES	$\Pi_{nm}^{REES,G}$	$\Pi_{nm}^{REES,LNG}$

Finally, note that in Section 5.3.4, we generally refer to these new indicators with the symbols Π_{nm}^G and Π_{nm}^{LNG} . In Section 5.5, we show the effects of the application of all these indicators that we have constructed.

5.3.3 Model assumptions

In this section, we illustrate the main assumptions that characterize our model formulation.

- **Producing/supplying countries.** In the considered gas supply chain, supplying countries produce natural gas and/or LNG and can decide to sell them directly to mid-streamers with LTCs or on the spot market. We assume that gas is extracted from sites directly owned by producers. The extraction process is supposed to be lossless. Therefore, the amount of gas extracted corresponds to the one produced. As indicated in the nomenclature, the variables x_{nmt}^G and x_{nmt}^{LNG} identify the amount of natural gas x_{nmt}^G and LNG x_{nmt}^{LNG} that the production country n sells to mid-streamer m in the low- and in the high-demand periods, respectively. We assume that these quantities as well as their relative prices depend on time to account for possible renegotiation or updates of the contract price in short term (see Franza (2014)). The quantity of natural gas sold on the spot market x_{nt}^{SpotG} depends on the production country n and on the time period t only. This assumption reflects the fact that the producer participates to the spot market by submitting bids but without knowing who will be the buyer. Finally, the quantity of LNG sold on spot basis $x_{nmt}^{SpotLNG}$ refers to uncommitted ships that are already arrived in the destination country and can be traded in the short term between the supplying country n and the mid-streamer m .
- **Destination country.** It is assumed that there is just one destination country that corresponds to Italy (see Section 5.4).
- **Mid-streamers.** We assume that all mid-streamers operate in the considered destination country and are located in one citygate. As already indicated, mid-streamers can decide to either enter in gas/LNG LTCs with supply countries or exchange gas/buy LNG on the corresponding spot markets. They select the supply countries from which importing natural gas and LNG on the basis of their external supply risks. As final step, they sell gas to the consumption sectors that we classify into power generation, industrial sector, and residential/commercial. In addition, we assume that mid-streamers operate the storage sites and can take advantage of seasonal arbitrage by buying and injecting gas into storage in the

low-demand season (summer) and then selling it to consumers in the high-demand season (winter). Parallel to the modeling of the producers' variables, we assume that the variables y_{nmt}^G and y_{nmt}^{LNG} respectively identify the amount of natural gas y_{nmt}^G and LNG y_{nmt}^{LNG} that the mid-streamer m buy from producing country n with LTCs in the two time periods. We further assume that mid-streamer can both buy and sell natural gas on the spot market. These mid-streamer's actions are identified by the variables y_{mt}^{SpotG} and q_{mt}^{SpotG} , respectively. Note that these variables depend on the mid-streamer m on the time period t only. This assumption reflects the fact that the mid-streamer participates with bids/offers to the spot market without knowing who will be the counterparts. Finally, the quantity of LNG purchased on spot basis $y_{nmt}^{SpotLNG}$ refers to uncommitted ships that can be traded in the short term between the n supplying and the m importing countries. As already explained, we do not model the mid-streamer's opportunity to re-sell the LNG acquired on the spot basis.

- **Consumers.** We assume that the gas demands of three consumer groups (industry, power generation, and residential/commercial) are endogenously determined through inverse demand functions. The variable d_{st} denotes the quantity of natural gas required by consumer s in time t .
- **Time framework and LTCs.** We consider a time framework of one year. Since LTCs have a duration of 20-25 years, we take a different approach to study the aforementioned “no flexibility” and the “flexibility” cases. In particular, in the case of “no flexibility”, we assume that, in the considered year, mid-streamers have to respect all the LTCs that they have already stipulated, but they can also decide to sign new LTCs. On the contrary, in the “flexibility” case, mid-streamers can ask for a re-negotiation of the existing contracts and therefore are not obliged to buy all the quantity of gas or LNG defined by the existing contracts.
- **Gas/LNG volumes and prices and degrees of mid-streamers' flexibility.** All gas/LNG volumes and prices are endogenously determined in the model. The volumes are defined through the supplying countries and mid-streamers' variables indicated above, while the LTCs and the spot prices correspond to the values of the dual variables associated with the relative gas and LNG balance constraints. Note that we impose some limits on the gas and LNG volumes contracted with LTCs. These limits vary according to the flexibility assumptions considered for mid-streamers. More precisely, we take as reference the annual volumes of gas and LNG that mid-streamers have contracted with already existing LTCs (these annual

volumes are input data) and in the “no flexibility” case we assume that these amounts establish lower bounds on the total quantities of gas and LNG that mid-streamers have to buy with LTCs in the year. This reflects the idea that mid-streamers have to respect all the LTCs that they have already stipulated, but they can also decide to sign new LTCs in the considered year. In contrast, to respect the re-negotiation assumption that characterizes the “flexibility” case, the amounts of already stipulated gas and LNG LTCs are used as upper bounds on the volumes of gas and LNG that mid-streamers can purchase.

- **Pipeline and tanker transportation limits.** For sake of simplicity, we assume that the tanker used to transfer LNG have no capacity limits. A similar assumption has been adopted for the gas pipelines. Pipeline gas and LNG transportation costs are included as exogenous charges in the model. Moreover, we assume that in the destination country there are some entry points for pipeline gas. Each of this entry point has a specific location in the destination country and collects the gas coming from a subset of supplying countries. We use the incidence matrix Γ_{fn} to define the link between supplying country and entry point. These entry points are characterized by a limited capacity that we model through a constraint.
- **External supply risk.** The external supply risk is considered as an externality that we internalize in our optimization problem. In general, the insecurity of supply could intervene at the time the prices are negotiated. As explained above, LTCs and spot prices for both gas and LNG are endogenously determined in our model. However, we assume that only gas and LNG traded with LTCs are exposed to the external supply risk. This is due to the fact that spot LNG refers to gas that is already landed in the destination country and therefore it is not risky. Similarly, the spot gas is freely traded on a short notice at the hub and therefore we suppose the external supply risk does not apply to its exchange. The inclusion of the risk indicators in our model is conducted as follows: we first construct and compute the values of the indicators presented in Section 5.3.2 on the basis of the ex-ante structure external supply of the considered destination country (Italy). We then use these indicators to weight the quantities of gas y_{nmt}^G and LNG y_{nmt}^{LNG} that appear in the respective LTC balance constraints (see model formulation in Section 5.3.4). We introduce the external supply risk indicators in the balances of the gas and LNG LTCs because these constraints not only represent the agreement between supplying countries and mid-streamers on the exchanged volumes but also are those that affect the prices of these LTCs. In this way, we can evaluate whether the external supply risk can modify the gas/LNG contract prices and volumes.

5.3.4 Optimization model for a natural gas supply chain with external supply risk

Finding the equilibrium state of the described natural gas chain with endogenous final consumers' demand consists in solving a social welfare optimization problem subject to some technical constraints. We first describe the cost functions that are included in our model formulation.

Supplying countries gain from selling gas but they face some costs in the gas production process. Let C_{nt} denote the production costs incurred by producer n . This is a continuous and convex function that depends on the total quantity of natural gas X_{nt}^G that supplying country n extracts in time period t . In particular, one has:

$$C_{nt} = C_{nt}(X_{nt}^G) \quad \forall n, \forall t \quad (5.14)$$

It could happen that a part of the produced gas is then sold as LNG by some of the supplying countries n . These countries face additional liquefaction costs that are denoted with the function LC_{nt} . This function is assumed to be continuous and convex and depends on the total amount of LNG X_{nt}^{LNG} that country n supplies in time period t :

$$LC_{nt} = LC_{nt}(X_{nt}^{LNG}) \quad \forall n, \forall t \quad (5.15)$$

On the other side, mid-streamers who import LNG in the destination (importing) country face the related regasification costs. These costs are represented by a continuous and convex function, denoted as RC_{mt} , that depends on the total quantity of LNG Y_{mt}^{LNG} that mid-streamer m regasifies in time period t . In particular, one has:

$$RC_{mt} = RC_{mt}(Y_{mt}^{LNG}) \quad \forall m, \forall t \quad (5.16)$$

In our gas supply chain, we further assume that mid-streamers manage the storage systems. They withdraw gas in the high-demand period $t = 2$ at zero cost,³ while the gas injection is operated in the low-demand period $t = 1$ with a cost that we denote with I_{m1} . This is assumed to be a continuous and convex function that depends on the quantity of natural gas i_{m1} that mid-streamer m injects in the storage site in $t = 1$. In particular, one has:

$$I_{m1} = I_{m1}(i_{m1}) \quad \forall m \quad (5.17)$$

³This assumption is taken from Egging et al. (2008).

In the following, we first describe the optimization model for the “no flexibility” case (Section 5.3.4) and we then illustrate the modifications that are introduced for modeling the “flexibility” assumption (Section 5.3.4).

“No flexibility” case

Finding the equilibrium state of the described natural gas chain with endogenous final consumers’ demand consists in solving the social welfare optimization problem (5.18)-(5.38) presented below. The objective function (5.18) corresponds to the (annual) social welfare that is given by the difference between the final consumers’ willingness to pay $\int_0^{d_{st}} P_{st}(\xi) d\xi$ and all the costs respectively faced by supplying countries and mid-streamers. In particular, supplying countries pay the production, the LNG liquefaction, the pipeline and the cargo transportation costs; on the other side mid-streamers bear the regasification, the storage injection, and the distribution charges. The objective function (5.18) is subject to several constraints as detailed in the following.

$$\begin{aligned}
 \text{Max} \quad & \sum_{s=1}^S \sum_{t=1}^T \theta_t \int_0^{d_{st}} P_{st}(\xi) d\xi - \sum_{n=1}^N \sum_{t=1}^T \theta_t \cdot [C_{nt}(X_{nt}^G) + LC_{nt}(X_{nt}^{LNG})] \\
 & - \sum_{n=1}^N \sum_{t=1}^T \theta_t \cdot \left[\sum_{m=1}^M ptc_{nm}^G \cdot x_{nmt}^G + \sum_{m=1}^M stc_{nm}^{LNG} \cdot x_{nmt}^{LNG} \right] + \\
 & - \sum_{n=1}^N \sum_{t=1}^T \theta_t \cdot \left[ptc_n^{SpotG} \cdot x_{nt}^{SpotG} + \sum_{m=1}^M stc_{nm}^{LNG} \cdot x_{nmt}^{SpotLNG} \right] + \\
 & - \sum_{m=1}^M \sum_{t=1}^T \theta_t \cdot [RC_{mt}(Y_{mt}^{LNG})] - \sum_{m=1}^M \theta_1 \cdot I_{m1}(i_{m1}) - \sum_{s=1}^S \sum_{t=1}^T \theta_t \cdot [dc_{ms} \cdot z_{mst}]
 \end{aligned} \tag{5.18}$$

subject to

$$\bar{X}_n - X_{nt}^G \geq 0 \quad \forall n, \forall t \quad (\bar{\gamma}_{nt}) \tag{5.19}$$

$$X_{nt}^G - \left(\sum_{m=1}^M x_{nmt}^G + x_{nt}^{SpotG} + X_{nt}^{LNG} \right) = 0 \quad \forall n, \forall t \quad (\gamma_{nt}) \tag{5.20}$$

$$(1 - \alpha_n) \cdot X_{nt}^{LNG} - \left(\sum_{m=1}^M x_{nmt}^{LNG} + \sum_{m=1}^M x_{nmt}^{SpotLNG} \right) = 0 \quad \forall n, \forall t \quad (\delta_{nt}) \quad (5.21)$$

$$\bar{L}_n - (1 - \alpha_n) \cdot X_{nt}^{LNG} \geq 0 \quad \forall n, \forall t \quad (\bar{\delta}_{nt}) \quad (5.22)$$

$$\bar{R}_m - (1 - \beta_m) \cdot Y_{mt}^{LNG} \geq 0 \quad \forall m, \forall t \quad (\bar{\eta}_{mt}) \quad (5.23)$$

$$(1 - \beta_m) \cdot Y_{mt}^{LNG} - \left(\sum_{n=1}^N y_{nmt}^{LNG} + \sum_{n=1}^N y_{nmt}^{SpotLNG} \right) = 0 \quad \forall m, \forall t \quad (\eta_{mt}) \quad (5.24)$$

$$\begin{aligned} & \sum_{n=1}^N y_{nmt}^G + \sum_{n=1}^N y_{nmt}^{LNG} + y_{mt}^{SpotG} + \sum_{n=1}^N y_{nmt}^{SpotLNG} + \\ & - i_{mt} \geq \sum_{s=1}^S z_{mst} + q_{mt}^{SpotG} \quad \forall m, t = 1 \quad (\lambda_{mt}) \end{aligned} \quad (5.25)$$

$$\begin{aligned} & \sum_{n=1}^N y_{nmt}^G + \sum_{n=1}^N y_{nmt}^{LNG} + y_{mt}^{SpotG} + \sum_{n=1}^N y_{nmt}^{SpotLNG} + \\ & + w_{mt} \geq \sum_{s=1}^S z_{mst} + q_{mt}^{SpotG} \quad \forall m, t = 2 \quad (\lambda_{mt}) \end{aligned} \quad (5.26)$$

$$i_{m1} \cdot \theta_1 - w_{m2} \cdot \theta_2 \geq 0 \quad \forall m \quad (\mu_m) \quad (5.27)$$

$$\bar{I}_m - i_{m1} \geq 0 \quad \forall m \quad (\nu_m) \quad (5.28)$$

$$\bar{W}_m - w_{m2} \geq 0 \quad \forall m \quad (\sigma_m) \quad (5.29)$$

$$WG_m - \theta_2 \cdot w_{m2} \geq 0 \quad \forall m \quad (\phi_m) \quad (5.30)$$

$$\sum_t \theta_t \cdot y_{nmt}^G - \tau_{nm} \geq 0 \quad \forall n, \forall m \quad (\psi_{mn}^G) \quad (5.31)$$

$$\sum_t \theta_t \cdot y_{nmt}^{LNG} - \xi_{nm} \geq 0 \quad \forall n, \forall m \quad (\psi_{mn}^{LNG}) \quad (5.32)$$

$$\Upsilon_{ft} - \left(\sum_{n=1}^N \sum_{m=1}^M \Gamma_{fn} \cdot y_{nmt}^G + \sum_{n=1}^N \Gamma_{fn} \cdot x_{nt}^{SpotG} \right) \geq 0 \quad \forall f, \forall t \quad (\kappa_{ft}) \quad (5.33)$$

$$x_{nmt}^G - \Pi_{nm}^G \cdot y_{nmt}^G = 0 \quad \forall n, \forall m, \forall t \quad (p_{nmt}^G) \quad (5.34)$$

$$x_{nmt}^{LNG} - \Pi_{nm}^{LNG} \cdot y_{nmt}^{LNG} = 0 \quad \forall n, \forall m, \forall t \quad (p_{nmt}^{LNG}) \quad (5.35)$$

$$\sum_{n=1}^N x_{nt}^{SpotG} + q_{mt}^{SpotG} - \sum_{m=1}^M y_{mt}^{SpotG} = 0 \quad \forall t \quad (p_t^{SpotG}) \quad (5.36)$$

$$x_{nmt}^{SpotLNG} - y_{nmt}^{SpotLNG} = 0 \quad \forall n, \forall m, \forall t \quad (p_{nmt}^{SpotLNG}) \quad (5.37)$$

$$\sum_m z_{mst} - d_{st} = 0 \quad \forall s, \forall t \quad (p_{st}) \quad (5.38)$$

Constraints (5.19)-(5.22) identify the supply countries' activities. In particular, constraint (5.19) imposes an upper bound on the total quantity of gas that producers can extract (X_{nt}^G). As stated by the constraint (5.20), the gas produced by supplying countries can be left in the gaseous form, and then sold either with LTCs (see variable x_{nmt}^G) or on the spot market (see variable x_{nt}^{SpotG}), or can be transformed in LNG (see variable X_{nt}^{LNG}). The variable X_{nt}^{LNG} represents the total amount of LNG produced by the supply country n that is then sold either with LTCs (x_{nmt}^{LNG}) or on spot ($x_{nmt}^{SpotLNG}$), as indicated by constraint (5.21). On the other side, constraint (5.22) imposes capacity limits on the liquefaction process. Note that the variable X_{nt}^{LNG} in constraints (5.21) and (5.22) is multiplied by the factor $(1 - \alpha_n)$ to account for the gas loss α_n that accrues during the liquefaction phase.

Constraints (5.23)-(5.32) refer to mid-streamers. In particular, mid-streamers buy gas and LNG from supplying countries with LTCs, or buy/sell gas on the spot market, or buy spot LNG. Mid-

streamers regasify the total amount of LNG purchased (Y_{mt}^{LNG}) taking into account the capacity of their technologies as indicated in constraint (5.23). On the other side, constraint (5.24) explains that Y_{mt}^{LNG} accounts for the LNG that mid-streamers purchase both with LTCs (y_{nmt}^{LNG}) and on spot ($y_{nmt}^{SpotLNG}$). Note that both in (5.23) and (5.24) the variable Y_{mt}^{LNG} is multiplied by the factor $(1 - \beta_m)$ to consider the losses of the regasification process. Constraints (5.25) and (5.26) define the balances among the quantities of gas managed by mid-streamer in the low ($t = 1$) and in the high-demand ($t = 2$) periods, respectively. More precisely, constraint (5.25) enforces that the total amount of gas purchased by the mid-streamer minus the gas injected in the storage site has to be greater or equal to the amount of gas sold to final consumers (z_{mst}) and on the spot market. In contrast, constraint (5.26) imposes that total amount of gas purchased by the mid-streamer plus the gas withdrawn from the storage site has to be greater or equal to the quantity of gas sold to final consumers (z_{mst}) and on the spot market. Constraints (5.27)-(5.30) regulate the storage process. In particular, (5.27) enforces that amount of gas injected in the storage site has to be greater than the quantity withdrawn. Constraints (5.28) and (5.29) respectively define the injection and the withdrawal capacity limits, and, finally, (5.30) imposes the working gas limit throughout all withdrawal periods. On the other side, constraints (5.31) and (5.32) are used to model the “no flexibility” assumption described above. In particular, they respectively impose that the yearly amounts of gas and LNG that mid-streamers have to buy through LTCs have to be greater or equal to the volumes established (for that year) in already existing contracts. Such a constraint formulation allows mid-streamers not only to accomplish the volume TOP clause of the LTCs into which they have already entered, but also to possibly negotiate new contracts. We explain in Section 5.3.4 how these constraints are modified to model the mid-streamers’ flexible behaviour.

Constraint (5.33) enforces the capacity limit of the entry points located in the destination country, while constraints (5.34), (5.35), (5.36) and (5.37) are the balances for gas and LNG respectively traded with LTCs and exchanged on a spot basis. Note that, as mentioned above, constraints (5.34) and (5.35) also include the external supply risk indicator as a weight of the quantities of gas and LNG purchased by the mid-streamers. Constraint (5.38) imposes the balance between the total quantity of gas sold by mid-streamers (z_{mst}) and demanded by consumers (d_{st}).

Finally, In order to detect the behaviour of the different players involved in the natural gas supply chain we consider complementarity formulation of this optimization problem:

$$0 \leq -\gamma_{nt} + \frac{\partial C_{nt}(X_{nt}^G)}{\partial X_{nt}^G} + \bar{\gamma}_{nt} \perp X_{nt}^G \geq 0 \quad \forall n, \forall t \quad (5.39)$$

$$0 \leq -(1 - \alpha_n) \cdot \delta_{nt} + (1 - \alpha_n) \cdot \bar{\delta}_{nt} + \gamma_{nt} + \frac{\partial LC_{nt}(X_{nt}^{LNG})}{\partial X_{nt}^{LNG}} \perp X_{nt}^{LNG} \geq 0 \quad \forall n, \forall t \quad (5.40)$$

$$0 \leq -p_{nmt}^G + pt c_{nm}^G + \gamma_{nt} \perp x_{nmt}^G \geq 0 \quad \forall n, \forall m, \forall t \quad (5.41)$$

$$0 \leq -p_t^{SpotG} + \gamma_{nt} + pt c_n^{SpotG} + \sum_{f=1}^F \Gamma_{fn} \cdot \kappa_{ft} \perp x_{nt}^{SpotG} \geq 0 \quad \forall n, \forall t \quad (5.42)$$

$$0 \leq -p_{nmt}^{LNG} + st c_{nm}^{LNG} + \delta_{nt} \perp x_{nmt}^{LNG} \geq 0 \quad \forall n, \forall m, \forall t \quad (5.43)$$

$$0 \leq -p_{nmt}^{SpotLNG} + st c_{nm}^{LNG} + \delta_{nt} \perp x_{nmt}^{SpotLNG} \geq 0 \quad \forall n, \forall m, \forall t \quad (5.44)$$

$$0 \leq -p_{st} + dc_{ms} + \lambda_{mt} \perp z_{mst} \geq 0 \quad \forall m, \forall s, \forall t \quad (5.45)$$

$$0 \leq -(1 - \beta_m) \cdot \eta_{mt} + \frac{\partial RC_m(Y_{mt}^{LNG})}{\partial Y_{mt}^{LNG}} + (1 - \beta_m) \cdot \bar{\eta}_{mt} \perp Y_{mt}^{LNG} \geq 0 \quad \forall m, \forall t \quad (5.46)$$

$$0 \leq -\lambda_{mt} - \psi_{nm}^G + \Pi_{nm}^G \cdot p_{nmt}^G + \sum_{f=1}^F \Gamma_{fn} \cdot \kappa_{ft} \perp y_{nmt}^G \geq 0 \quad \forall n, \forall m, \forall t \quad (5.47)$$

$$0 \leq -\lambda_{mt} + p_t^{SpotG} \perp y_{mt}^{SpotG} \geq 0 \quad \forall m, \forall t \quad (5.48)$$

$$0 \leq -\lambda_{mt} - \psi_{nm}^{LNG} + \Pi_{nm}^{LNG} \cdot p_{nmt}^{LNG} + \eta_m \perp y_{nmt}^{LNG} \geq 0 \quad \forall n, \forall m, \forall t \quad (5.49)$$

$$0 \leq -\lambda_{mt} + p_{nmt}^{LNG} + \eta_m \perp y_{nmt}^{SpotLNG} \geq 0 \quad \forall m, \forall t \quad (5.50)$$

$$0 \leq -p_t^{SpotG} + \lambda_{mt} \perp q_{mt}^{SpotG} \geq 0 \quad \forall m, \forall t \quad (5.51)$$

$$0 \leq -\mu_m + \frac{\partial I_{m1}(i_{m1})}{\partial i_{m1}} + \nu_m + \lambda_{m1} \perp i_{m1} \geq 0 \quad \forall m, t = 1 \quad (5.52)$$

$$0 \leq -\lambda_{m2} + \mu_m + \sigma_m + \phi_m \perp w_{m2} \geq 0 \quad \forall m, \forall t = 2 \quad (5.53)$$

$$0 \leq \bar{X}_n - X_{nt}^G \perp \bar{\gamma}_{nt} \geq 0 \quad \forall n, \forall t \quad (5.54)$$

$$0 \leq \bar{L}_n - (1 - \alpha_n) \cdot X_{nt}^{LNG} \perp \bar{\delta}_{nt} \geq 0 \quad \forall n, \forall t \quad (5.55)$$

$$0 \leq \bar{R}_m - (1 - \beta_m) \cdot Y_{mt}^{LNG} \perp \bar{\eta}_{mt} \geq 0 \quad \forall m, \forall t \quad (5.56)$$

$$0 \leq \sum_{n=1}^N y_{nmt}^G + \sum_{n=1}^N y_{nmt}^{LNG} + y_{mt}^{SpotG} + \sum_{n=1}^N y_{nmt}^{SpotLNG} +$$

$$-i_{mt} - \sum_{s=1}^S z_{mst} - q_{mt}^{SpotG} \perp \lambda_{mt} \geq 0 \quad \forall m, t = 1 \quad (5.57)$$

$$0 \leq \sum_{n=1}^N y_{nmt}^G + \sum_{n=1}^N y_{nmt}^{LNG} + y_{mt}^{SpotG} + \sum_{n=1}^N y_{nmt}^{SpotLNG} + w_{mt} + \sum_{s=1}^S z_{mst} - q_{mt}^{SpotG} \perp \lambda_{mt} \geq 0 \quad \forall m, t = 2 \quad (5.58)$$

$$0 \leq i_{m1} - w_{m2} \perp \mu_m \geq 0 \quad \forall m \quad (5.59)$$

$$0 \leq \bar{I}_m - i_{m1} \perp \nu_m \geq 0 \quad \forall m \quad (5.60)$$

$$0 \leq \bar{W}_m - w_{m2} \perp \sigma_m \geq 0 \quad \forall m \quad (5.61)$$

$$0 \leq WG_m - \theta_2 \cdot w_{m2} \perp \phi_m \geq 0 \quad \forall m \quad (5.62)$$

$$0 \leq \sum_t \theta_t \cdot y_{nmt}^G - \tau_{nm} \perp \psi_{mn}^G \geq 0 \quad \forall n, \forall m \quad (5.63)$$

$$0 \leq \sum_t \theta_t \cdot y_{nmt}^{LNG} - \xi_{nm} \perp \psi_{mn}^{LNG} \geq 0 \quad \forall n, \forall m \quad (5.64)$$

$$0 \leq \Upsilon_{ft} - \left(\sum_{n=1}^N \sum_{m=1}^M \Gamma_{fn} \cdot y_{nmt}^G + \sum_{n=1}^N \Gamma_{fn} \cdot x_{nt}^{SpotG} \right) \perp \kappa_{ft} \geq 0 \quad \forall t \quad (5.65)$$

$$0 \leq p_{st} - a_{st} + b_{st} \cdot d_{st} \perp d_{st} \geq 0 \quad \forall s, \forall t \quad (5.66)$$

$$X_{nt}^G - \left(\sum_{m=1}^M x_{nmt}^G + x_{nt}^{SpotG} + X_{nt}^{LNG} \right) = 0 \quad \forall n, \forall t \quad (\gamma_{nt} : \text{free}) \quad (5.67)$$

$$(1 - \alpha_n) \cdot X_{nt}^{LNG} - \left(\sum_{m=1}^M x_{nmt}^{LNG} + \sum_{m=1}^M x_{nmt}^{SpotLNG} \right) = 0 \quad \forall n, \forall t \quad (\delta_{nt} : \text{free}) \quad (5.68)$$

$$(1 - \beta_m) \cdot Y_{mt}^{LNG} - \left(\sum_{n=1}^N y_{nmt}^{LNG} + \sum_{n=1}^N y_{nmt}^{SpotLNG} \right) = 0 \quad \forall m, \forall t \quad (\eta_{mt} : \text{free}) \quad (5.69)$$

$$x_{nmt}^G - \Pi_{nm}^G \cdot y_{nmt}^G = 0 \quad \forall n, \forall m, \forall t \quad (p_{nmt}^G : \text{free}) \quad (5.70)$$

$$\sum_{n=1}^N x_{nt}^{SpotG} + q_{mt}^{SpotG} - \sum_{m=1}^M y_{nmt}^{SpotG} = 0 \quad \forall t \quad (p_t^{SpotG} : \text{free}) \quad (5.71)$$

$$x_{mnt}^{LNG} - \Pi_{nm}^{LNG} \cdot y_{mnt}^{LNG} = 0 \quad \forall n, \forall m, \forall t \quad (p_{mnt}^{LNG} : \text{free}) \quad (5.72)$$

$$x_{nmt}^{SpotLNG} - y_{nmt}^{SpotLNG} = 0 \quad \forall n, \forall m, \forall t \quad (p_{nmt}^{SpotLNG} : \text{free}) \quad (5.73)$$

$$\sum_m z_{mst} - d_{st} = 0 \quad \forall s, \forall t \quad (p_{st} : \text{free}) \quad (5.74)$$

“Flexibility” case

The formulation of the welfare optimization problem under the “flexibility” assumption is identical to that presented in Section 5.3.4 with the exception of the constraints regulating the volumes of

gas and LNG volumes that mid-streamers have to buy with LTCs. In particular, this change regards constraints (5.31) and (5.32) that from lower bounds on the gas and LNG volumes purchased with LTCs become upper bounds. These are expressed as follows:

$$\tau_{nm} - \sum_t \theta_t \cdot y_{nmt}^G \geq 0 \quad \forall n, \forall m \quad (\psi_{mn}^G) \quad (5.75)$$

$$\xi_{nm} - \sum_t \theta_t \cdot y_{nmt}^{LNG} \geq 0 \quad \forall n, \forall m \quad (\psi_{mn}^{LNG}) \quad (5.76)$$

From a mathematical point of view, this constraint modification implies some small changes in the KKT formulation of the optimization problem. In particular, conditions (5.47), (5.49), (5.63), and (5.64) are respectively substituted with the following ones:

$$0 \leq -\lambda_{mt} + \psi_{nm}^G + \Pi_{nm}^G \cdot p_{nmt}^G + \sum_{f=1}^F \Gamma_{fn} \cdot \kappa_{ft} \perp y_{nmt}^G \geq 0 \quad \forall n, \forall m, \forall t \quad (5.77)$$

$$0 \leq -\lambda_{mt} + \psi_{nm}^{LNG} + \Pi_{nm}^{LNG} \cdot p_{nmt}^{LNG} + \eta_m \perp y_{nmt}^{LNG} \geq 0 \quad \forall n, \forall m, \forall t \quad (5.78)$$

$$0 \leq \tau_{nm} - \sum_t \theta_t \cdot y_{nmt}^G \perp \psi_{mn}^G \geq 0 \quad \forall n, \forall m \quad (5.79)$$

$$0 \leq \xi_{nm} - \sum_t \theta_t \cdot y_{nmt}^{LNG} \perp \psi_{mn}^{LNG} \geq 0 \quad \forall n, \forall m \quad (5.80)$$

More precisely, since constraints (5.75) and (5.76) impose upper bounds on the primal variables y_{nmt}^G and y_{nmt}^{LNG} , the associated dual variables ψ_{mn}^G and ψ_{mn}^{LNG} enter with a positive sign in the KKT conditions (5.77) and (5.78) of these primal variables. The reverse happens in the complementarity formulation of the optimization problem under the “no flexibility” assumption. Since constraints (5.31) and (5.32) define lower bounds on variables y_{nmt}^G and y_{nmt}^{LNG} the associated dual variables ψ_{mn}^G and ψ_{mn}^{LNG} enter with a negative sign in the KKT conditions (5.47) and (5.49) of these primal variables. Finally, all the other KKT conditions are as indicated in Section 5.3.4. From an economical point of view, the illustrated constraint modification allows us to describe a more flexible behaviour of the mid-streamers which, in this case, have the possibility to ask for a LTC re-negotiation or to avoid the respect of the contract TOP clause for a short period.

5.4 Case study

Our case study is based on the Italian gas market that we consider as the gas destination area where mid-streamers operate. We select this market for two reasons: first, it is one of the three largest gas markets in Europe together with the UK and Germany; second, it mainly relies on natural gas imports to cover its demand since the national production is very low. According to the annual report of the Italian Authority (see *Autorità per l'Energia Elettrica il Gas e il Sistema Idrico* (2016)⁴) the 90.6% of the national gas demand in 2015 was satisfied with imports from Russia, Algeria, Libya, the Netherlands, Qatar, and Norway. The main companies (mid-streamers) operating in Italy are ENI, Edison, and Enel with a market share of respectively 53.8%, 21.2%, and 11.2% (see *Autorità per l'Energia Elettrica il Gas e il Sistema Idrico* (2016)). In 2015, these companies bought natural gas from Russia, Algeria, Libya, the Netherlands, and Norway; while LNG was imported by Qatar and Algeria (see *Autorità per l'Energia Elettrica il Gas e il Sistema Idrico* (2016)). In general, residential/commercial is the gas consumers' sector with the highest demand, followed by power generation, and industrial sector (see also *Autorità per l'Energia Elettrica il Gas e il Sistema Idrico* (2016)). The Italian natural gas and LNG imports are mainly delivered via long-term contracts, even though gas can also be traded on the Italian spot market "Punto di Scambio Virtuale (PSV)" that was created in 2003 (see Honoré (2013) for a description of the liberalization process of the Italian gas market and the establishment of the PSV).

Considering this framework, our analysis refers to 2015 data and it is based on the following assumptions:

- **Destination country:** Italy.
- **Supplying countries:** Russia (RU), Algeria (AL), Libya (LIB), the Netherlands (NL), Qatar (QT), and Norway (NW). Since in the last years the Italian gas production is progressively reducing (see *Autorità per l'Energia Elettrica il Gas e il Sistema Idrico* (2016)), we do not account for Italy among the producing and supplying countries.
- **Natural gas origin:** Russia, Algeria, Libya, the Netherlands, and Norway. For modeling the mid-streamers' flexibility assumptions, we consider the LTCs that Italy has established with these countries in the last years and are still active in 2015.

⁴This annual report is in Italian and refer to 2015 data. An English version is available at http://www.autorita.energia.it/allegati/relaz_ann/15/annual_report2015.pdf but refers to 2014 data.

- **LNG origin:** Algeria and Qatar (see BP (2016); International Group of Liquefied Natural Gas Importers (2018)). Similarly to gas, we consider the LNG LTCs that Italy has established with these two countries in the last years and are still active in 2015 to model the mid-streamers' flexibility assumptions.
- **Mid-streamers:** We assume that there is just one mid-streamer, since we do not dispose of detailed data for all the companies operating in the Italian gas market. This representative mid-streamer can buy both natural gas and LNG through LTCs or can trade them on the respective spot markets.

As already explained, we account for a time span of one year subdivided into two time periods with different demand levels. The high-demand period is assumed to have a duration of 151 days and comprehends the months from November to March included; the low-demand period lasts 214 days and covers the remaining months. Since, in our analysis, we consider a representative day per each period, all quantities are expressed in mcm/day while prices and costs are in €/cm.

The production (extraction) capacity data of the aforementioned supplying countries are taken from Egging et al. (2008).⁵ These data are not recent but considering that the gas reserve to production (R/P) ratio computed at worldwide level has been almost unchanged in the last twenty years,⁶ we consider these capacity data as a reasonable proxy. The liquefaction capacity data related to the supplying countries are taken from International Group of Liquefied Natural Gas Importers (2018) and refer to 2015.

No capacity limits are imposed on the natural gas transports via pipelines between supplying countries and Italy or on the LNG cargos. However, we consider the capacity of the entry points located at the borders of the Italian network that enforces restrictions on the amount of natural gas imported via pipelines. Italy has five entry points for pipelines that are Mazara del Vallo, Gela, Tarvisio, Passo Gries, and Gorizia. In Mazara del Vallo, natural gas is imported from Algeria thanks to the connection with the pipeline Transmed/Enrico Mattei; the natural gas from Libya enters in Gela through the connection with the pipeline Greenstream; Gorizia and Tarvisio receive gas from Russia through the TAG pipeline, and, finally, Pass Gries gets gas from the Netherlands and Norway respectively via the Trans-European pipeline and the Transitgas. Considering this information about the gas provenience at the different entry points, we are able to limit the imports between

⁵See Table 14 at page 2410 of Egging et al. (2008).

⁶See BP at <http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/natural-gas/natural-gas-reserves.html>

Italy and the supplying countries (see constraint (5.33) in our formulation). The 2015 capacity data of these entry points are provided by Snam (2016). For what concerns LNG, the capacity of the regasification plants implicitly limits the Italian LNG imports. There are three regasification plants in Italy that are located in Rovigo, Livorno, and Panigaglia. In our case study, we consider just one regasification plant whose capacity is obtained by aggregating those of these two plants. The respective data are taken from International Group of Liquefied Natural Gas Importers (2018) and refer to 2015.

The natural gas production costs faced by supplying countries are defined as linear function of the following type:

$$C_{nt} = c_n^G \cdot X_{nt}^G \quad \forall n, \forall t$$

where the parameter c_n^G has been estimated taking as reference Egging et al. (2008). In particular, the value of c_n^G for the considered supplying countries is obtained by multiplying the marginal production costs reported in the column “Max mag” of Table 14 in Egging et al. (2008) by 1.9 that is the average for 2015 of the “Euro Area Gross Domestic Product Chained 2010 Prices YoY”.⁷

The liquefaction costs incurred by supplying countries are determined by the following quadratic function:

$$LC_{nt} = lc1_n^{LNG} \cdot X_{nt}^{LNG} + lc2_n^{LNG} \cdot (X_{nt}^{LNG})^2 \quad \forall n, \forall t$$

where the terms $lc1_n^{LNG}$ and $lc2_n^{LNG}$, as for the production costs, have been estimated taking as reference the data and the approach proposed by Egging et al. (2008). More precisely, the values of $lc1_n^{LNG}$ and $lc2_n^{LNG}$ respectively correspond to the α and β parameters reported in Table 15 of in Egging et al. (2008) multiplied by 1.9. i.e. the 2015 GDP.

For the Italian mid-streamer’s regasification costs, we consider the following quadratic function:

$$RC_t = rc1^{LNG} \left(\sum_{n=1}^N Y_{nt}^{LNG} \right) + rc2^{LNG} \left(\sum_{n=1}^N Y_{nt}^{LNG} \right)^2 \quad \forall t$$

where the terms $rc1_n^{LNG}$ and $rc2_n^{LNG}$ have been evaluated using the data available in Egging et al.

⁷Gross domestic product (GDP) measures the final market value of all goods and services produced within a country. It is the most frequently used indicator of economic activity. The GDP by expenditure approach measures total final expenditures (at purchasers’ prices), including exports less imports. This concept is adjusted for inflation. For our simulation, we GDP data from Bloomberg (ticker: EUGDEMU). See <https://www.bloomberg.com/quote/EUGNEMUY:IND>

(2008). More precisely, the values of $rc1_n^{LNG}$ and $rc2_n^{LNG}$ respectively correspond to the α and β parameters reported in Table 16 of in Egging et al. (2008) multiplied by 1.9. i.e. the 2015 GDP.

The mid-streamer also controls and manages the storage site. We suppose that there is just one storage site whose capacity and injection rate are obtained by aggregating the capacities and injection rates of all the storage sites available in Italy. As in Egging and Gabriel (2006), we impose that the injection rate is equal to the peak output rates, while the extraction capacity is assumed to be twice the injection capacity. The data related to working gas and storage rates refer to 2015 and are taken from International Energy Agency (2016a) and from the Stogit website.⁸ Injection costs are defined through the following linear function:

$$I_1 = ic \cdot i_1$$

where the parameter ic is the unitary injection cost whose value has been taken from Autorità per l'Energia Elettrica il Gas e il Sistema Idrico (2016).

Supplying countries also face the natural gas and LNG transportation costs. The costs of transfer natural gas through pipelines are taken from NERA economic consulting (2014). In particular, we have taken the 2018 data in Figure 77 at page 132 of Appendix A and we have adjusted them with the pipeline cost adders provided in Figure 87 at page 139 of Appendix A for 2018. We have finally transformed all cost data from \$/Mcf to €/cm. On the other side, for evaluating the LNG transportation costs we follow the same approach adopted for production, liquefaction, and regasification costs. We consider a cost of 0.005€/cm/1000 sea miles that we have taken from Egging et al. (2008) and we have then multiplied it by 1.9. The sea miles are deduced from International Group of Liquefied Natural Gas Importers (2018). On the other side, the distribution costs faced by the mid-streamer are taken from Autorità per l'Energia Elettrica il Gas e il Sistema Idrico (2016).

The data needed for the computation of the external supply risk indicators refer to 2015 and are taken from BP (2016); International Group of Liquefied Natural Gas Importers (2018); Autorità per l'Energia Elettrica il Gas e il Sistema Idrico (2016). As in Le Coq and Paltseva (2009), we consider the political risk rating for 2015 published by the PRS Group.⁹ This political risk measure assigns countries a rate whose values are between 1 and 100 with the following reasoning: the highest the rate, the lowest the political risk associated. In all indicators that we construct, we consider the complementary of this PRS risk, namely r_n is set in the following way:

⁸See <http://www.stogit.it/en/about-us/where-you-can-find-us/storage-sites.html>

⁹See <https://www.prsgroup.com/category/risk-index>

$$r_n = 100 - \text{PRS}_{\text{Risk}}$$

Recall that the indicators $\Pi_{nm}^{REES,G}$ and $\Pi_{nm}^{REES,LNG}$ also account for the distance (d_{nm}) between the supplying and the destination countries and the fungibility of natural gas (F_{nm}^G) and the LNG (F_{nm}^{LNG}). Following Le Coq and Paltseva (2009), we set $F_{nm}^G = 1$ and $F_{nm}^{LNG} = 2$, and the transport risk is identified by the following parameter d_{nm} :

$$d_{nm} = \begin{cases} 1, & \text{distance between capitals} < 1700 \text{ Km} \\ 2, & 1700 \text{ Km} \leq \text{distance between capitals} < 2500 \text{ Km} \\ 3, & 2500 \text{ Km} \leq \text{distance between capitals} < 3300 \text{ Km} \\ 4, & \text{distance between capitals} \geq 3300 \text{ Km} \end{cases} \quad (5.81)$$

The external risk indicators that we obtain from our computations are reported in Table 5.2.

Table 5.2: External supply risk indicators

	$\Pi_{nm}^{HHI,G}$	$\Pi_{nm}^{SWN2,G}$	$\Pi_{nm}^{GES,G}$	$\Pi_{nm}^{REES,G}$
NW	1.598	2.877	0.102	0.115
NL	1.174	3.974	0.113	0.126
RU	18.779	19.140	11.564	6.464
AL	1.508	8.784	0.300	0.335
LIB	1.377	15.494	0.495	0.553
	$\Pi_{nm}^{HHI,LNG}$	$\Pi_{nm}^{SWN2,LNG}$	$\Pi_{nm}^{GES,LNG}$	$\Pi_{nm}^{REES,LNG}$
AL	0.111	3.515	0.001	0.003
QT	93.444	0.360	0.973	0.799

In order to model the different degrees of mid-streamer's flexibility, we need the data of the annual volumes of gas and LNG that have been regulated by LTCs in 2015. We recall that Italy has natural gas LTCs with Russia, Algeria, the Netherlands, Norway, and Libya; while LNG LTCs are with Algeria and Qatar. The considered annual data are reported in Table 5.3 and are taken from Cedigaz¹⁰

We consider the following inverse demand functions to model the final consumers' gas demand:

¹⁰See <http://www.cedigaz.org/products/natural-gas-database.aspx>

$$p_{st} = a_{st} - b_{st} \cdot d_{st} \quad \forall s, \forall t$$

Parameters a_{st} and b_{st} have been estimated using an elasticity value of -0.1 for all consumer groups, the amount of gas demanded by Italian industry, power sector, and residential/commercial and the average prices that they paid for purchasing gas in 2015. These data are taken from Autorità per l'Energia Elettrica il Gas e il Sistema Idrico (2016).

Table 5.3: Volumes of gas and LNG regulated by LTCs between Italy and supplying countries in 2015 (annual value)

	Gas	LNG
NW	5,450	-
NL	9,200	-
RU	35,000	-
AL	27,380	1,840
LIB	5,610	-
QT	-	6,400
Tot	82,640	8,240

Finally, we recall that the solution of the optimization problems presented in Sections 5.3.4 and 5.3.4 is found by implementing their complementarity conditions. These KKT conditions are run in the GAMS modeling environment, using PATH as solver (see Dirkse and Ferris (1995)).

5.5 Results

To analyze the impacts of external supply risk on the mid-streamer imports' choices and we consider the following assumptions

1. Mid-streamer's behaviour:

- The mid-streamer has *at least* to buy an amount of gas and LNG as defined in already existing LTCs ("No FLEX" case in the following);
- The mid-streamer has *at most* to buy an amount of gas and LNG as defined in already existing LTCs ("FLEX" case in the following).

2. External supply risk:

- No external supply risk is considered in model (“NO Risk” case in the following);
- All external supply risk indicators that we constructed are considered in the model as explained in the previous sections (“Risk” case in the following).

In the following we first analyze the impacts of the application of different degrees of mid-streamers’ flexibility by comparing the “No FLEX” and “FLEX” without considering any of the external supply risk indicators (see Section 5.5.1). The effects of gas/LNG volumes and prices with risk are then presented in Section 5.5.2. Note that in this section we consider the risk indicators imposed in both the “No FLEX” and the “FLEX” cases. Note that, in any of the simulations that we have implemented, the mid-streamer re-sells gas on the spot market. Finally, we assume that the risk regards both the supply of gas and LNG with LTCs at the same time.

5.5.1 No external supply risk

We first consider the effects of the mid-streamer’s flexibility on the annual amount of gas purchased. Fig. 5.2a illustrate the yearly volume of natural gas, both in gaseous and liquefied forms, that the mid-streamer exchanges in the “No FLEX” and the “FLEX”) cases, assuming that the external supply risk is not considered (“NO Risk”).

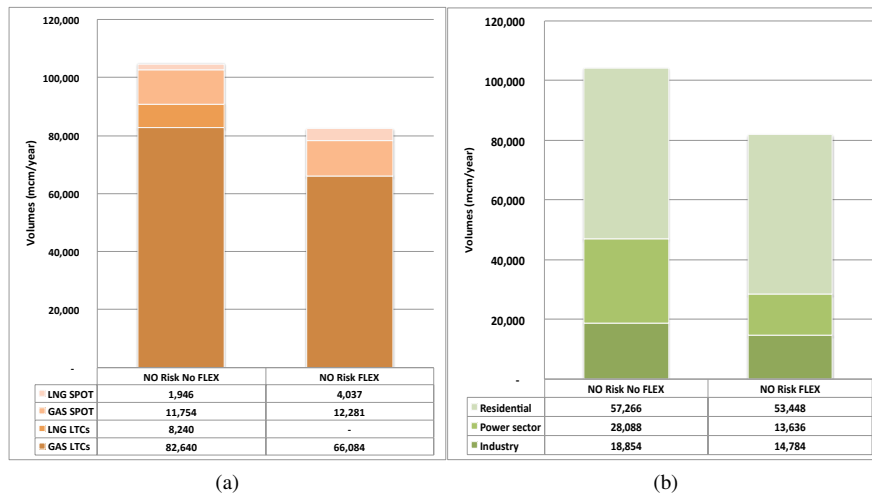


Figure 5.2: Yearly volume of gas exchanged per type and yearly demand per consumer group (mcm/year)

The mid-streamer accomplishes all gas and LNG LTCs that it has with all the supplying countries (compare the values reported in Figure 2a with those reported in Table 5.3), while this does not happen under the “FLEX” assumption leading to a gas volume drop of 21%. In particular, the mid-streamer reduces by 20% the annual volume of gas purchased with LTCs and does no longer buy LNG with LTCs. Even though it increases respectively by 4% and 107% the amount of gas and LNG acquired on spot markets at yearly basis, these spot purchases do not suffice to compensate the reduction in the LTC volumes. This has a negative impact on final consumers that see their gas availability reduced in the “FLEX” case. In particular, the mid-streamer still guarantee almost the entire gas supply to the residential/commercial sector (the drop is “only” by 7%), but it decreases the supply to the power companies and industry respectively by 51% and 22% with respect to the “No FLEX” case (see Fig. 5.2b).

Fig. 5.3 and 5.4 provide more details on the daily amount of gas and LNG exchanged between supplying countries and mid-streamers in the low- and in the high-demand periods under the “No FLEX” and “FLEX” assumptions. In general, we observe that between the flexibility and not flexibility cases, there is no a drastic variation of the mix of supplying countries. What the mid-streamer significantly modifies is the amount of gas/LNG purchased in the two time periods (shift of volumes exchanged in low- and high-demand periods) and the selection of the physical status of the gas traded (shift between the quantities of gas and LNG purchased, especially for the supply from Algeria). The set of supplying countries from which the mid-streamer receives gas/LNG is not subject to many modifications. Considering LTCs (Fig 5.3), the main changes in the gas supply mix between the “No FLEX” and “FLEX” cases are as follows: under the flexibility assumption, Norway does no longer provide gas to Italy in any period; Libya still procures Italy with gas but only in the low-demand period; and, finally, the total amount of gas provided by Algeria reduces, especially in the low-demand period. Note that this unsold Algerian gas is then transformed in LNG and sold on the spot market (see Fig. 5.4. For what concerns LNG, as already indicated above, the difference between the “No FLEX” and “FLEX” cases is that the volume TOP clause of the Algerian and Qatari LTCs is not respected under the flexibility assumption because the mid-streamer prefers to buy spot LNG. On the other side, both in the “No FLEX” and “FLEX” cases, Libya and the Netherlands supply spot gas to Italy, even though in different quantities. Moreover, in the high-demand period of the “FLEX” case Russia appears as an additional supplier.

The comparison of the “No FLEX” and “FLEX” results under the assumption of absence of external supply risk shows that the mid-streamer changes its supply choices when it has the possibility to

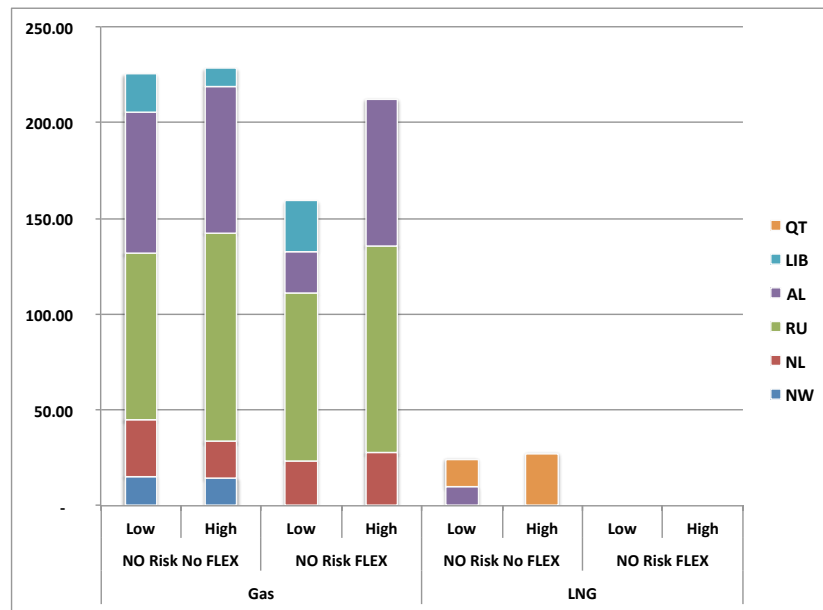


Figure 5.3: Natural gas and LNG bought through LTCs (mcm/day)

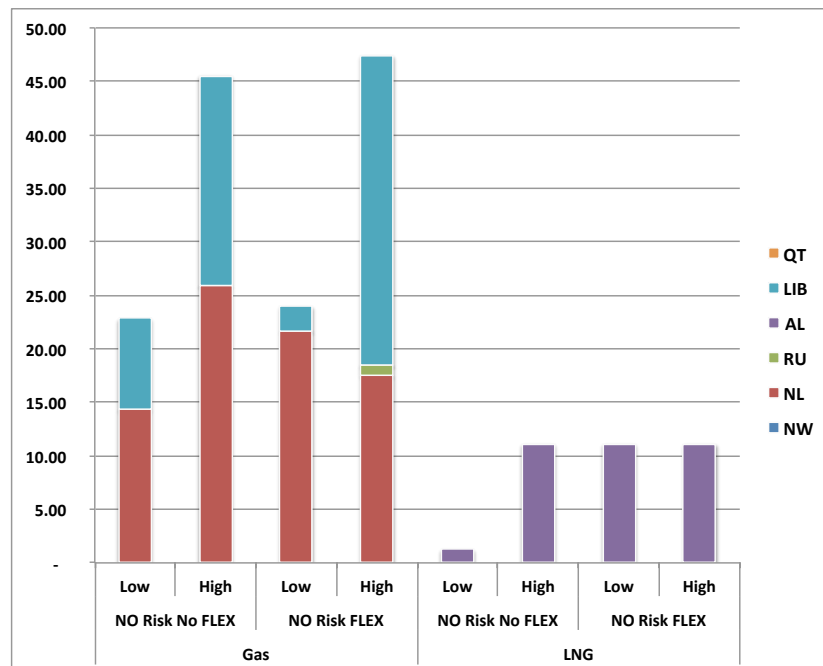


Figure 5.4: Natural gas and LNG bought on the spot market (mcm/day)

do that. The profit analysis can help understanding this change of strategies in the mid-streamer's behaviour. Table 5.4 reports supplying countries and the mid-streamer's profits, the final gas con-

Table 5.4: Supplying countries and mid-streamer's profits, final consumers' surplus, and social welfare under the "NO Risk" assumption (€/year)

€/cm		NO Risk No FLEX	NO Risk FLEX
Supplying countries	Revenues	51,021	38,342
	Production costs	14,745	11,171
	Transport costs	36,057	27,188
	Net profits	218	-17
Mid-streamer	Revenues	50,247	52,550
	Distribution costs	11,326	9,895
	Purchase costs	51,021	38,342
	Regasification costs	149	53
	Injection costs	161	150
	Net profits	-12,410	4,110
Final consumers	Industry's surplus	21,820	13,962
	Power Sector 's surplus	14,454	4,899
	Residenatial's surplus	182,618	159,711
	Total surplus	218,892	178,572
Social welfare		206,700	182,664

sumers' surplus, and the social welfare in the "No FLEX" and "FLEX" cases. In the "No FLEX" case, the mid-streamer is forced to satisfy the LTCs that it has already contracted. This guarantees the remuneration of the supplying countries which gain from their activities, but it causes a significant loss for the mid-streamer because its cost of purchasing gas and LNG ("Purchase costs") is higher than the revenues it obtains from selling gas to final consumers (compare Tables 5.5, 5.6, and 5.7 that report the LTC prices, the spot prices, and the prices charged to final consumers, respectively).¹¹ The situation changes in the "FLEX" case where the mid-streamer is not obliged to fully satisfy all the contracts. The outcome is that the mid-streamer modifies its supply mix in such a way it avoids negative profits. However, the mid-streamer's strategy of reducing its purchase of gas and LNG with LTCs compared to the "No FLEX" case is not beneficial for the whole gas supply chain because it implies not only a reduction of gas availability for final consumers with a consequent decrease of their surplus, but also a drop of supplying countries' profits that become

¹¹Note that, in the "No FLEX" case, the weighted average gas and LNG prices computed over the involved supplying countries and paid by the mid-streamer are: 0.50 €/cm (gas LTCs), 0.66 €/cm (LNG LTCs), 0.32 €/cm (gas spot), and 0.50 €/cm (LNG spot). In contrast, the weighted average price that the mid-streamer applies to final consumers is 0.48 €/cm).

negative (see Table 5.4). The final result is a 12% reduction of the social welfare. We recall that the remuneration of supplying countries is important because these guarantee the maintenance of the existing infrastructures and the investments in new ones, whose costs are, in general, very high. Considering these results, we can say that LTCs are necessary to maintain the stability and to guarantee the security of supply, even though this may incur in losses for the mid-streamers. We finally reports in Tables 5.5 and 5.6 the prices for natural gas and LNG respectively sold with LTCs and on the spot market obtained in the “No FLEX” and the “FLEX” cases. Note that the label “n.s.” in Tables 5.5 and 5.6 stands for “not sold”. We recall that the gas spot price is not differentiated per supplier because we assume that there is just one market where all participants submit their offers and bids. This market sets the price for spot gas on a daily basis. The prices of LTCs for gas and LNG are higher than those defined on the spot market both under the “No FLEX” and the “FLEX” assumptions. This reflects the differences between these two pricing systems and explains the reason why mid-streamers have asked for the re-negotiation of the LTCs. Indeed, higher LTCs prices guarantee returns to supplying countries, but also they may lead to possible losses for mid-streamers. Moreover, as expected, the spot prices in the low-demand period are lower than in those in high-demand period because of the different consumption levels in the two time frameworks.

Table 5.5: Prices of gas and LNG LTCs under the “NO Risk” assumption (€/cm)

		NW	NL	RU	AL	LIB	QT
Gas LTCs	No FLEX	0.54	0.38	0.51	0.54	0.41	
	FLEX	n.s.	0.38	0.51	0.55	0.41	
LNG LTCs	No FLEX				0.51		0.72
	FLEX				n.s.		n.s.

Table 5.6: Prices spot of gas and LNG under the “NO Risk” assumption (€/cm)

			Low	High
Gas Spot	No FLEX		0.31	0.33
	FLEX		0.25	0.32
LNG Spot	No FLEX	AL	0.49	0.50
		QT	n.s.	n.s.
	FLEX	AL	0.31	0.38
		QT	n.s.	n.s.

Looking at Table 5.5, one can see that the gas prices of the LTCs with the Netherlands, Russia, and Libya are identical in the “No FLEX” and “FLEX” cases. In fact, the quantities of gas LTCs

exchanged between these countries and Italy are the same under the two flexibility assumptions. This explains the effects on the volume re-shuffle between low- and high-demand periods to which we assist in the “FLEX” case, as explained above. The unique exception is Algeria where the LTCs price in the “FLEX” case is higher than under the “No FLEX” assumption. This is the reason why, in the “FLEX” case, the mid-streamer decides to reduce the amount of gas purchased with LTCs from Algeria and buy this gas on the spot basis at lower prices (compare Algerian prices in Tables 5.5 and 5.6).

Considering Table 5.6, one can notice that the prices of the gas and LNG traded on the spot markets are lower in the “FLEX” case than in the “No FLEX” instance. This is the reason why the mid-streamer increase its purchases on spot.

Finally Table 5.7 reports the prices applied in the two time periods to final consumers in the “no FLEX” and the “FLEX” cases. Note that the prices that the mid-streamer is able to apply to final consumers under the “FLEX” assumption are higher than those imposed in the “no FLEX” case. This indeed allows mid-streamer to increase its revenues and compensate its costs in such a way it does not incur in a profit loss (see Table 5.4).

Table 5.7: Final consumers’ prices under the “NO Risk” assumption (€/cm)

	NO Risk No FLEX		NO Risk FLEX	
	Low	High	Low	High
Industry	0.39	0.48	0.57	0.60
Power sector	0.37	0.45	0.55	0.58
Residential/commercial	0.47	0.56	0.65	0.68

5.5.2 External supply risk

In this section, we analyze the effects of the application of the external supply risk in the “No FLEX” and “FLEX” cases. We recall that the risk is applied only on gas and LNG LTCs.

The main outcomes are as follows. The external supply risk does not affect the mid-streamer’s behaviour in the “No FLEX” case since the total volume of exchanged gas and LNG both via LTCs and on spot are as under the “NO Risk No FLEX” assumption (see Fig. 5.5) with the consequence that the total quantity of gas offered to final consumers remains unchanged (see Fig. 5.13). Moreover, no changes in the supplying country mix is encountered with respect to the not-flexible risk without

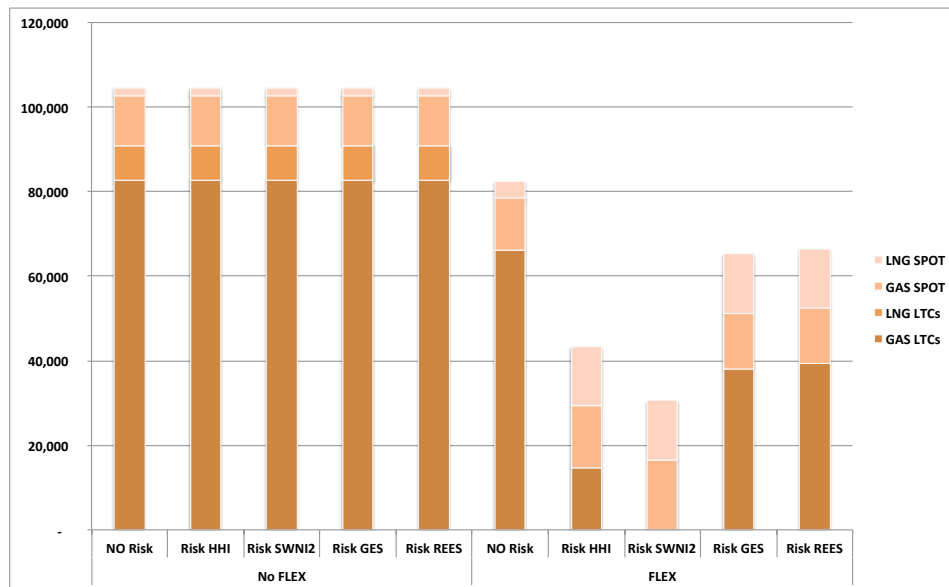


Figure 5.5: Yearly volume of gas exchanged per type (mcm/year)

risk. As in the “NO Risk No FLEX” case, the mid-streamer accomplishes the natural gas and LNG LTCs independently of the risk measure considered (see Fig. 5.6 and 5.7).

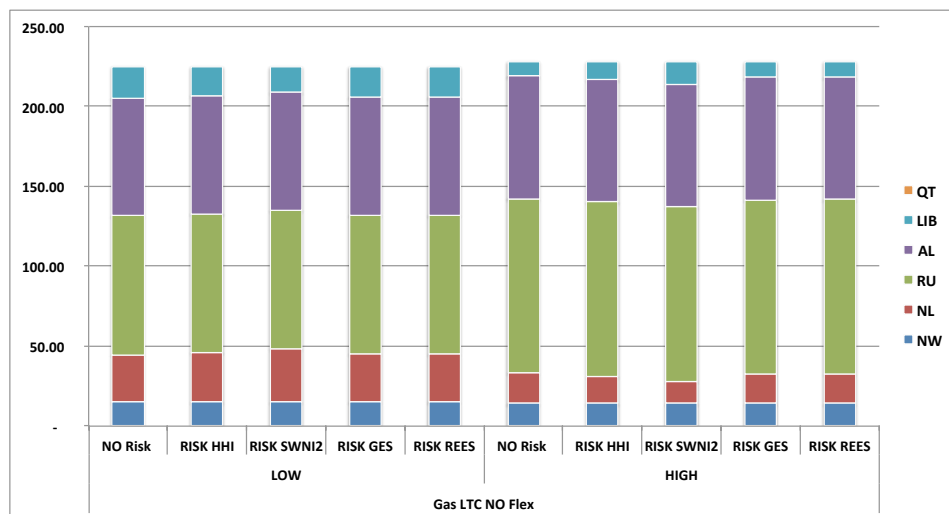


Figure 5.6: Volumes of gas exchanged via LTCs under the “Risk” and “No FLEX” assumptions (mcm/day)

This depends on the fact that the mid-streamer must satisfy the LTC volume clause and it has, at least, to buy the volumes of gas and LNG that it has already contracted. The risk does not either

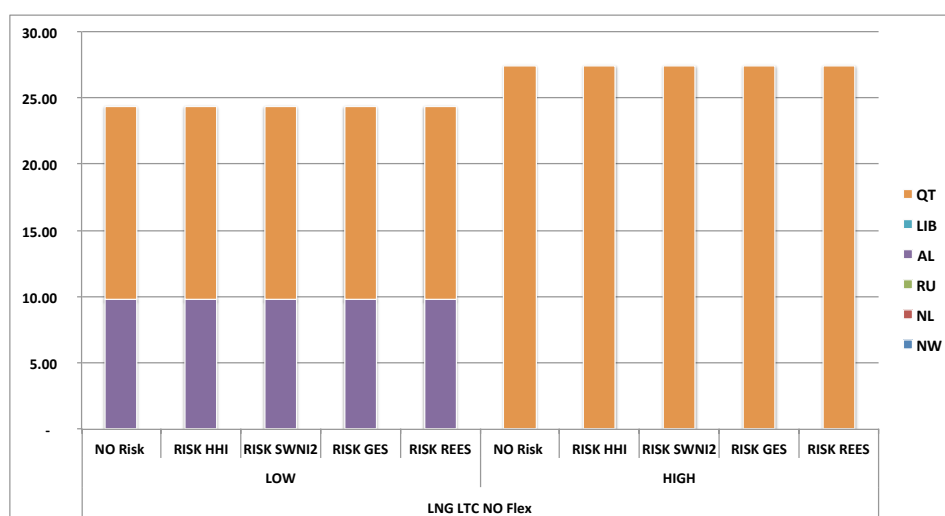


Figure 5.7: Volumes of LNG exchanged via LTCs under the “Risk” and “No FLEX” assumptions (mcm/day)

affect the spot supply, even though the mid-streamer does not have any constraints on the amount of gas and LNG that it has to buy on the spot markets (see Fig. 5.8 and 5.9).

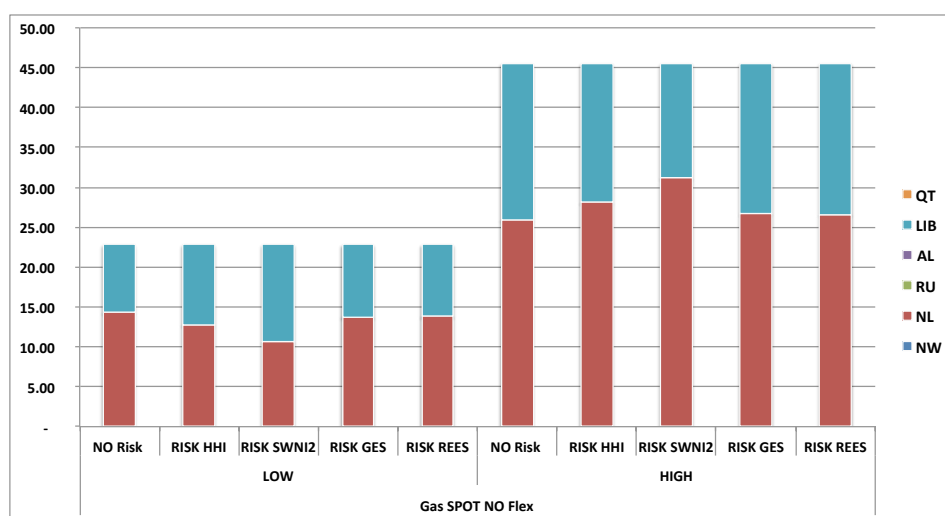


Figure 5.8: Volumes of spot gas exchanged under the “Risk” and “No FLEX” assumptions (mcm/day)

In other words, the external supply risk does not have significant impacts on the mid-streamer’s strategies when no flexibility is allowed. We register only some shifts between the quantities of the LTCs and spot gas purchased in the low- and in the high-demand period (see Fig. 5.6 and 5.8,

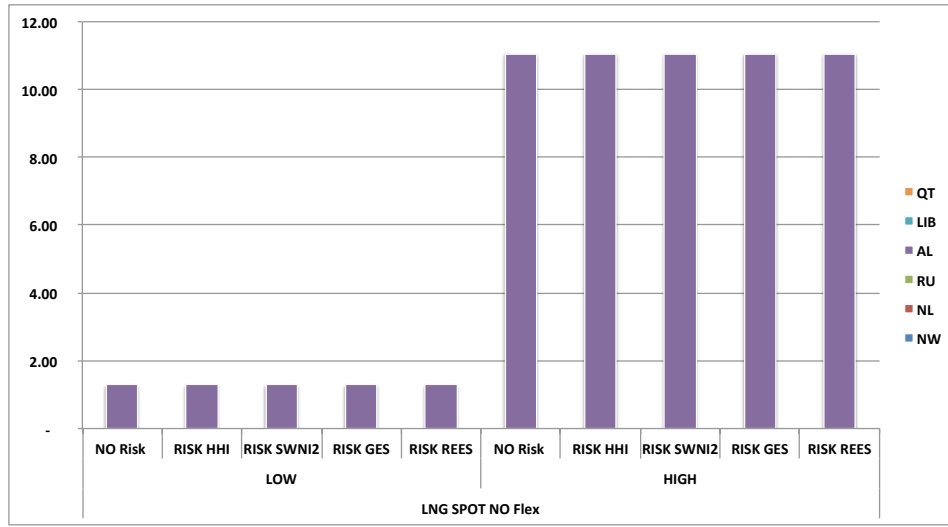


Figure 5.9: Volumes of natural gas exchanged via LTCs under the “Risk” and “No FLEX” assumptions (mcm/day)

respectively). This implies that the gas and LNG prices as well as those paid by the final consumers in all “Risk No Flex” cases are identical to those reported in Tables (5.5)–(5.7) for the corresponding “NO Risk No Flex” case, independently of the considered external supply risk indicator. The same holds for the social welfare of the gas supply chain and the profits/surplus of its player groups. The situation in the “Risk FLEX” cases remains in line with that described under the “NO Risk FLEX” assumption and, depending on the considered supply risk indicator, it is even more exacerbated. The mid-streamer’s actions in the “Risk FLEX” cases can be summarized with the following items (see Fig. 5.5 and Fig. 5.10-5.12):

1. The total amount of natural gas that is traded in the “Risk FLEX” cases is lower than under the “NO Risk FLEX” assumption. This is particular evident when the Π_{nm}^{HHI} and $\Pi_{nm}^{SWNI_2}$ indicators are applied. Globally, we register a significant drop of gas imports compared to the “No FLEX” cases;
2. In the “Risk FLEX” cases, the mid-streamer modifies the mix of supplying countries depending on the applied risk indicator;
3. The amount of gas that the mid-streamer exchanges with LTCs in the “Risk FLEX” cases is lower than the corresponding one in the “NO Risk FLEX” because the risk induces the mid-streamer to buy cheaper gas and LNG on the respective spot markets;

4. The volumes of gas and LNG that are exchanged on the spot markets increases under the “Risk FLEX” assumptions compared to the “NO Risk FLEX” (and also “NO Risk No FLEX”) cases, but this increase does not suffice to recover the amount of gas that is not bought with LTCs;
5. As in the “NO Risk FLEX” case, also under all “Risk FLEX” assumptions, the mid-streamer does not respect the volume TOP clause of the LNG LTCs; i.e. no LNG is exchanged with contracts.

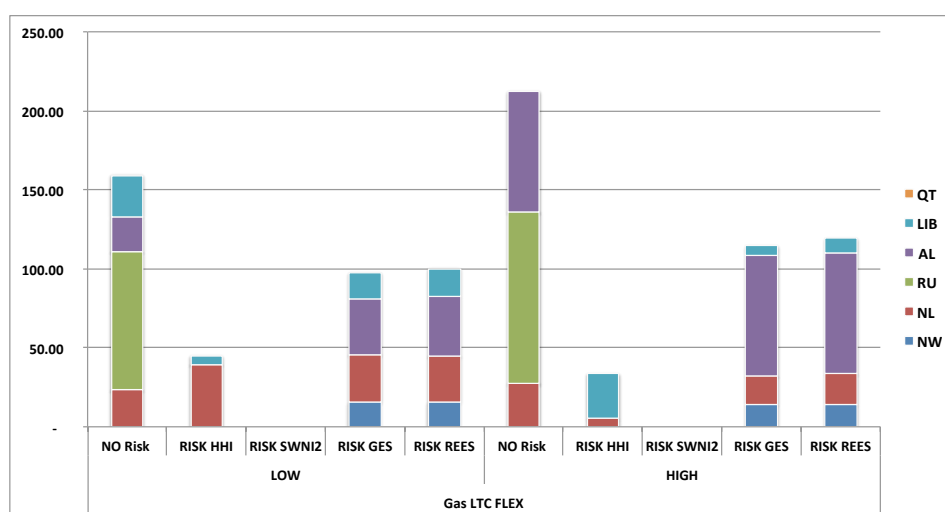


Figure 5.10: Volumes of natural gas exchanged via LTCs under the “Risk” and “FLEX” assumptions (mcm/day)

Table 5.8 provides a summary of the supplying countries that exchanges the different types of gas with Italy under the risk and flexibility assumptions. These are compared to the results obtained in the “NO Risk FLEX” case. For each case analyzed, “Yes” in the table cells indicates that there is a trade and the subsequent percentage reported in brackets corresponds to the weight in terms of volumes that the considered country has in the Italian supply mix. The symbol “-” means that no exchanges of gas or LNG are allowed between Italy and the producer, while “-No” means that Italy does not import from the specific country even though it has the possibility to do it. The provenience of spot gas and LNG is not affected by risk since the relative indicators are applied only on LTCs. As we already observed, there is an increase of the total volume of spot gas and LNG that is imported both from Algeria and Qatar, but this increase is not enough to restore the gas volume availability guaranteed by the respect of LTCs. The inclusion of risk has a significant impact on

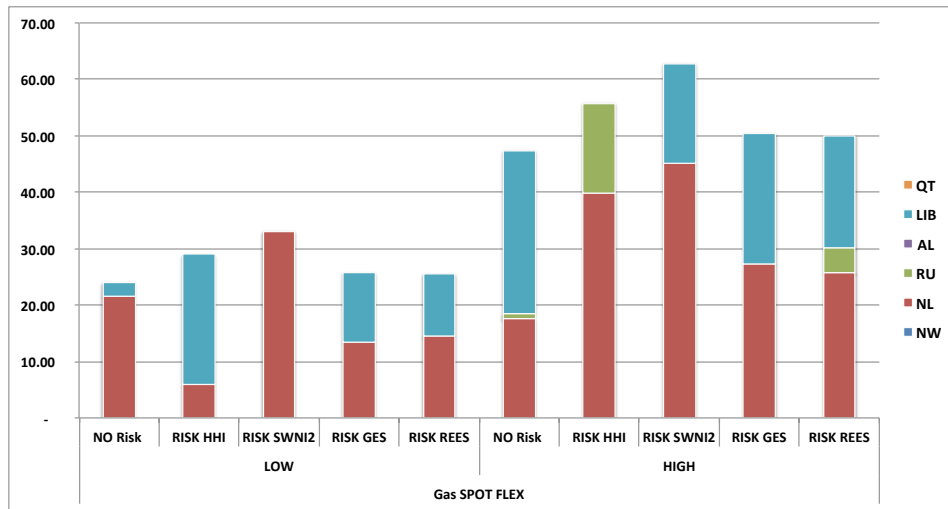


Figure 5.11: Volumes of spot gas exchanged under the “Risk” and “FLEX” assumptions (mcm/day)

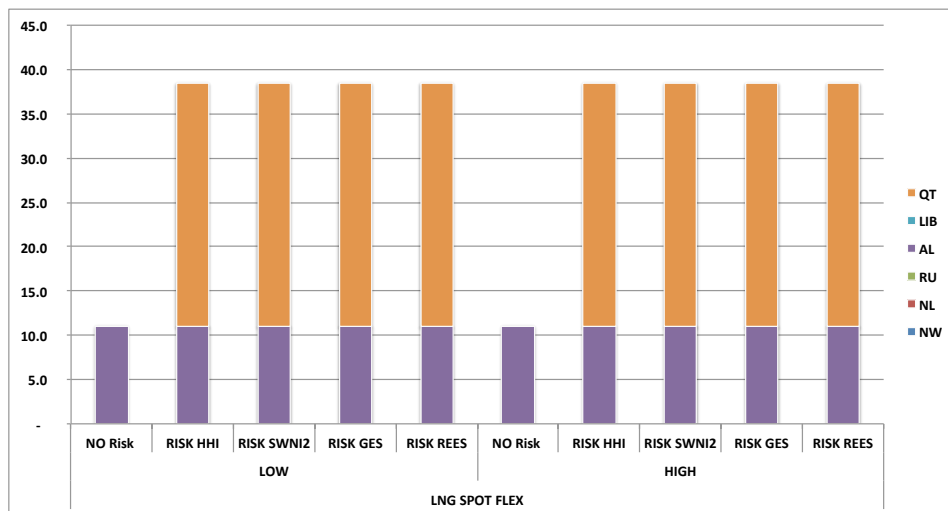


Figure 5.12: Volumes of spot LNG exchanged under the “Risk” and “FLEX” assumptions (mcm/day)

the external supply of gas with LTCs because the mid-streamer changes the provenience and the proportion of gas imported depending on the indicator. The principle driving the mid-streamer’s choice is considering not risky countries. For instance, under the Π_{nm}^{HHI} , the mid-streamer decides to fully respect the gas LTCs with the Netherlands and Libya because these are the two countries with the lowest risk level (see $\Pi_{nm}^{HHI,G}$ in Table 5.2). The same happens when applying the other risk

indicators. Note that the gas contracts with Russia are not honored because this is the producer with the highest risk in all considered indicators and, moreover, the mid-streamer does not accomplish to any gas LTCs under the $\Pi_{nm}^{SWNI_2}$ since the risk values associated with the different countries in this indicator are extremely high, much higher than those of the other indicators (see Table 5.2). On the other side, the mid-streamers does not buy LNG with contracts but prefers to resort to its purchase on spot, as it happens in the “NO Risk FLEX” case. This drop of natural gas import with LTCs has been already observed in Section 5.5.1 when describing the effect of the “NO Risk FLEX” case. The combination of the flexibility with the risk assumption exacerbates the phenomenon, implying a significant reduction of the total gas offered to final consumers in the “Risk FLEX” cases (see Fig. 5.13). Only the residential/commercial sector maintains a relative adequate supply level, even though lower than in the corresponding “Risk No FLEX” and “Risk FLEX” cases, whereas industries and power generators face a significant curtailment. These two sectors are not even supplied when the $\Pi_{nm}^{SWNI_2}$ risk applies.

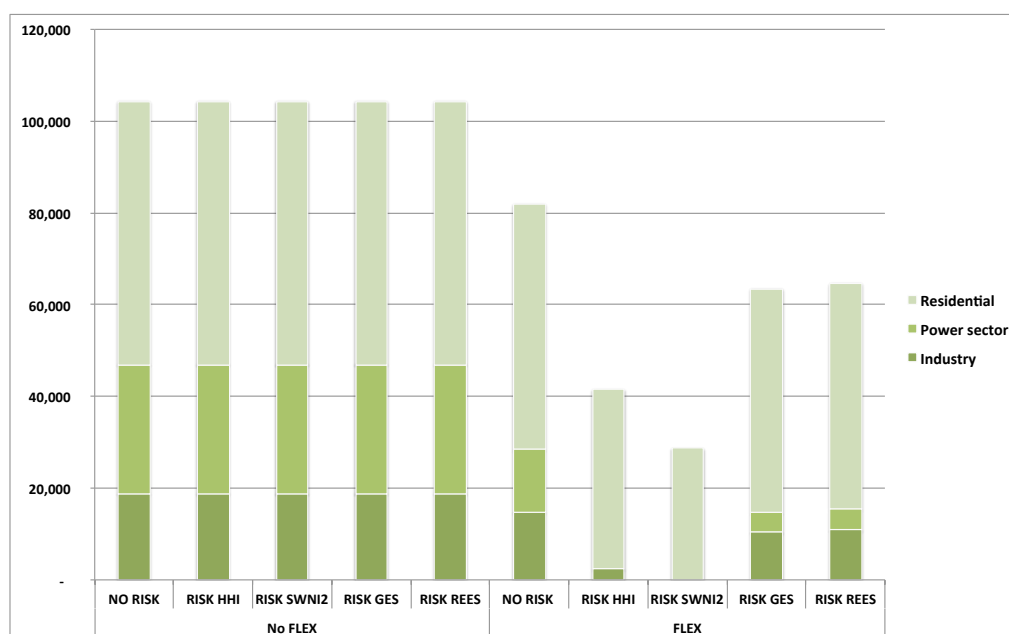


Figure 5.13: Yearly demand per consumer group (mcm/year)

The mid-streamer behaves in this way because it wants to limit its exposure to risk, but mostly because it desires to mitigate its possible profit losses. This issue has already been detected in Section 5.5.1 when discussing the “NO Risk FLEX” case.

Table 5.9 reports supplying countries and the mid-streamer’s profits, the final gas consumers’ sur-

plus, and the social welfare in the “NO Risk FLEX” and in the “Risk FLEX” cases. In all “FLEX” cases, the mid-streamer’s profits are positive thanks to the set of strategies that it adopts. Note that the significant curtailment of the expensive gas volumes traded with LTCs under the risk assumption allows the mid-streamer to increase its net profit compared to the “NO Risk FLEX” case. This is also due to the fact that the prices imposed to final consumers are higher than under the “NO Risk FLEX” assumption (see Table 5.11). This is particular evident when the $\Pi_{nm}^{SWNI_2}$ risk is implemented, since no gas/LNG is purchased with LTCs, but also when applying the Π_{nm}^{HHI} where the amount of gas/LNG bought on spot is proportionally higher than that imported with LTCs. This indeed has a negative impact on final consumers that see their surplus reducing with respect to the “NO Risk FLEX” assumption. In fact, the lowest consumers’ surplus is registered when the $\Pi_{nm}^{SWNI_2}$ risk is applied. The effects of the risk implementation on supplying countries’ profits vary according to the risk indicator analyzed. In particular, supplying countries globally face losses when the risk indicators Π_{nm}^{GES} and Π_{nm}^{REES} are considered, while they gain with the Π_{nm}^{HHI} and $\Pi_{nm}^{SWNI_2}$ measures. Note that these losses are higher than in the “NO Risk FLEX” case. In this latest case, it is true that their revenues are lower, but also the associated production and transportation costs are limited. This allows supplying countries to gain from the situation.

Table 5.10 reports the LTC prices for gas under the “NO Risk FLEX” and “Risk FLEX” assumptions. One can note that the risk leads to the a slight increase of these prices.

In conclusion, the “FLEX” strategy is more protecting for the mid-streamers, but it does not result to be beneficial for the whole supply chain that registers a reduction of the social welfare. This phenomenon becomes more extreme with the application of the external supply risk measures.

5.6 Conclusions

In this work, we analyze the security of the external supply of the Italian gas market that mainly relies on imports to satisfy its gas demand. In particular, we develop an optimization problem model that describes the equilibrium state of a natural gas supply chain where supplying countries, mid-streamers and consumers exchange natural gas and LNG both with LTCs and on the spot market. Mid-streamers who buy natural gas and LNG are assumed to be the market player mainly exposed to the external supply risk associated with the imports from foreign countries. In other words, mid-streamers define the amount of gas and LNG to be imported not only on the basis of the relative production and transportation costs, but also on the external supply risk associated with the countries from which the gas originates. The external supply risk is measured through indicators

that we construct starting from those already existing in the literature. These indicators are then inserted in the volume balance constraints of the gas and LNG LTCs.

In addition to the impact of the external supply risk, we analyze different degrees of mid-streamer's flexibility. In particular, we consider both a situation where the mid-streamer fully satisfies the LTCs quantity clause and a case where the mid-streamer behaves in a more flexible way and it is not obliged to fulfill the LTC volume requirements.

Our analysis shows that, if the mid-streamer have to comply with the LTCs quantity clause ("No FLEX" assumption), it does not significantly change its supplying choices even when the risk is considered. Under this assumption, the mid-streamer faces significant losses, while the supplying countries gain. In contrast, these mid-streamers' losses disappear when it is not obliged to fully satisfy the LTCs requirements ("FLEX" case) because it is able to modify its supply mix. In particular, compared to the "No FLEX" case, it reduces the amount of gas imported with LTCs because it is more expensive and increases the quantity of cheaper spot gas. In addition, it increases its imports from less risky countries when possible. However, this flexible mid-streamer's behaviour has several drawbacks compared to the "No FLEX" case: the suppliers can face losses because of the significant drop in their revenues; the total amount of gas and LNG purchased drops because the decrease of the quantity of gas imported via LTCs is not fully compensated by the increase of the spot gas or LNG. This also leads to a reduction of gas availability for final consumers. In particular, the mid-streamer still guarantee the gas supply to the residential/commercial sector, but it decreases the supply to the power companies and to industries with respect to the "No FLEX" case. Considering these results, it turns out that LTCs are necessary to maintain the stability and to guarantee the security of supply, even though this may incur in losses for the mid-streamers. The "FLEX" strategy is more protecting for the mid-streamers, but it does not result to be beneficial for the society. This phenomenon becomes more extreme with the application of the external supply risk measures.

Table 5.8: Supplying countries that exchange gas and LNG with Italy under the “Risk” and “FLEX” assumptions

		NW	NL	RU	AL	LIB	QT
Gas LTCs	NO Risk	No	Yes (14%)	Yes (54%)	Yes (25%)	Yes (8%)	-
	Risk Π_{nm}^{HHI}	No	Yes (62%)	No	No	Yes (38%)	-
	Risk $\Pi_{nm}^{SWNI_2}$	No	No	No	No	No	-
	Risk Π_{nm}^{GES}	Yes (15%)	Yes (24%)	No	Yes (50%)	Yes (11%)	-
	Risk Π_{nm}^{REES}	Yes (14%)	Yes (23%)	No	Yes (50%)	Yes (13%)	-
LNG LTCs	NO Risk	-	-	-	No	-	No
	Risk Π_{nm}^{HHI}	-	-	-	No	-	No
	Risk $\Pi_{nm}^{SWNI_2}$	-	-	-	No	-	No
	Risk Π_{nm}^{GES}	-	-	-	No	-	No
	Risk Π_{nm}^{REES}	-	-	-	No	-	No
Gas Spot	NO Risk	No	Yes (59%)	Yes (1%)	No	Yes (40%)	-
	Risk Π_{nm}^{HHI}	No	Yes (50%)	Yes (16%)	No	Yes (34%)	-
	Risk $\Pi_{nm}^{SWNI_2}$	No	Yes (84%)	No	No	Yes (16%)	-
	Risk Π_{nm}^{GES}	No	Yes (53%)	No	No	Yes (47%)	-
	Risk Π_{nm}^{REES}	No	Yes (54%)	Yes (5%)	No	Yes (41%)	-
LNG Spot	NO Risk	-	-	-	Yes (100%)	-	No
	Risk Π_{nm}^{HHI}	-	-	-	Yes (29%)	-	Yes (71%)
	Risk $\Pi_{nm}^{SWNI_2}$	-	-	-	Yes (29%)	-	Yes (71%)
	Risk Π_{nm}^{GES}	-	-	-	Yes (29%)	-	Yes (71%)
	Risk Π_{nm}^{REES}	-	-	-	Yes (29%)	-	Yes (71%)

Table 5.9: Supplying countries and mid-streamer's profits, final consumers' surplus, and social welfare under the "NO Risk" assumption (€/year)

FLEX						
		No Risk	Risk Π_{nm}^{HHI}	Risk $\Pi_{nm}^{SWNI_2}$	Risk Π_{nm}^{GES}	Risk Π_{nm}^{REES}
Supplying countries	Revenues	38,342	23,346	21,272	30,178	30,692
	Production costs	11,171	6,986	5,232	10,139	10,285
	Transportation costs	27,188	11,880	8,214	20,166	20,630
	Net Profits	-17	4,480	7,825	-128	-223
Mid-streamer	Revenues	52,550	49,877	45,647	52,481	52,464
	Distribution costs	9,895	6,247	4,471	8,472	8,579
	Purchase costs	38,342	23,346	21,272	30,178	30,692
	Regasification costs	53	185	185	185	185
	Injection costs	150	133	79	160	161
	Net Profits	4,110	19,966	19,640	13,486	12,846
Final consumers	Industry's surplus	13,962	961	-	7,436	7,934
	Power Sector's surplus	4,899	-	-	945	1,122
	Residenatials' surplus	159,711	86,322	47,453	133,285	135,427
	Total surplus	178,572	87,283	47,453	141,666	144,483
Social welfare		182,664	111,728	74,918	155,024	157,107

Table 5.10: Prices of gas LTCs under the "Risk FLEX" assumptions (€/cm)

FLEX						
	NW	NL	RU	AL	LIB	QT
NO Risk	n.s.	0.38	0.51	0.55	0.41	
Risk Π_{nm}^{HHI}	n.s.	0.39	n.s.	n.s.	0.43	
Risk $\Pi_{nm}^{SWNI_2}$	n.s.	n.s.	n.s.	n.s.	n.s.	
Risk Π_{nm}^{GES}	0.56	0.41	n.s.	0.56	0.43	
Risk Π_{nm}^{REES}	0.55	0.40	n.s.	0.56	0.42	

Table 5.11: Final consumers' prices under the "Risk FLEX" assumptions (€/cm)

FLEX										
	NO Risk FLEX		Risk Π_{nm}^{HHI}		Risk $\Pi_{nm}^{SWNI_2}$		Risk Π_{nm}^{GES}		Risk Π_{nm}^{REES}	
	Low	High	Low	High	Low	High	Low	High	Low	High
Industry	0.57	0.60	1.05	1.13	n.s.	n.s.	0.74	0.77	0.73	0.76
Power Sector	0.55	0.58	n.s.	n.s.	n.s.	n.s.	0.65	0.75	0.65	0.73
Residenatials	0.65	0.68	1.18	1.21	1.56	1.59	0.82	0.86	0.81	0.84

Chapter 6

Closure

6.1 Summary

This thesis can be seen as an attempt to study relevant issues related to the decarbonization of electricity and gas markets. Renewable energy generation is seen as one of the most effective ways to curb climate change, and intermittent sources like wind and solar are considered to be those with the highest potential. Anyway, renewable generation implies dealing with issues related to the variability and predictability of these sources, which represent a challenge for the power system. Possible solutions include integrating storage units to provide flexibility and use gas-fired plants as backup capacity. Because of the crucial role played by gas market in achieving the decarbonization targets, issues related to gas procurement in Europe are also analyzed. In particular, we investigate the re-negotiation of the LTCs required by European mid-streamers to make LTC gas prices more aligned with those resulting from the spot trading at the hubs. Moreover, we consider the issue of the risk of external supply, with a particular focus on the Italian gas market. The need of investigating this topic arises from the fact that most of the European Member States, and especially Italy, relies on imports from politically and economically unstable countries to cover their demand.

6.2 Conclusions

In this section, we briefly recall the main conclusions derived from the work developed in this thesis. Please refer to chapter related conclusions for additional details.

Policy mechanisms are needed to internalize the societal cost of carbon and address the trans-

formation of the electric power system required by the decarbonization process. Onshore wind represents the best option to reach immediate emission reduction. Anyway, if CSP achieves its projected cost decrease and efficiency gains, it can become an extremely competitive alternative. Moreover, the role of storage is fundamental in a renewable-dominated electric power system because it provides flexibility and allows for costs saving by serving electricity during peak hours. Electricity should also contribute to the decarbonization of the transport sector by facilitating the integration of electric vehicles. Benefits in terms of carbon emission reduction deriving from the EVs' usage are already tangible with the actual electricity generation mix, but these are expected to grow with the progressive increase of the renewable energy penetration expected in the years to come.

As indicated in the Energy Roadmap 2050, natural gas is the cheapest fossil fuel, and combined cycle power plants represent the most convenient electricity generation units in terms of both cost/efficiency ratio and emission rate. From a technological point of view, our analysis shows that the integration of the CCS mechanism in gas-fired power plants is a plausible option for the future in order to completely abate their carbon emissions. This is a general outcome that could be valid for all gas markets. However, considering our analysis of the European gas market, we can conclude that the LTCs need to be re-negotiated by also including, among the oil-based underlying assets, spot prices of gas traded at the hubs. Under this framework, introducing spot gas indexation might help mid-streamers' to compensate for possible losses caused by the accomplishment to quantity clause that anyway it is important to maintain the security of gas supply. The quantity clause is also a key issue in the analysis of the risk associated with external gas supply. Our analysis shows that, if the mid-streamer have to comply with the LTCs quantity clause, it does not significantly change its supplying choices even when the risk is considered. Under this assumption, the mid-streamer faces significant losses, while the supplying countries gain. In contrast, these mid-streamers' losses disappear when it is not obliged to fully satisfy the LTCs requirements because it is able to modify its supply mix.

6.3 Future Research

The models developed in the thesis project could be improved and further developed under different aspects. For this reason, we provide future research suggestions in the following:

- To extend the model proposed in Chapter 2 by developing a multi-stage stochastic optimiza-

tion planning investment model with decarbonization targets.

- To consider alternative clustering methods, such as hierarchical agglomerative clustering using Ward's method, which has been shown to perform well when dealing with a combination of renewable sources with atypical pattern (wind) and storage.
- To increase the number of operating conditions and scenarios and adopt a Bender's decomposition approach for the solution of the model investigated in the Chapters 2 and 3.
- To apply the models proposed in Chapter 2 to a realistic case study on the European electricity market and include pumped hydro storage as standalone storage unit.
- To consider smart charging strategies and vehicle-to grid systems in the planning of a SSEES as possible development of model investigated in Chapter 3.
- To develop the model formulated in Chapter 5 by introducing stochasticity in the demand by considering future consumption evolution by sector in the equilibrium problem on the gas supply chain.

List of Figures

Chapter 1

1.1	Electric Power System structure	4
1.2	Representation of the gas supply chain participant's and interactions	20

Chapter 2

2.1	28-Node Network Diagram	56
2.2	Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case	64
2.3	Expected Annual Operating and Investment Costs In Business-As-Usual Case and With Different Carbon-Emissions Reductions Relative to Business-As-Usual Case .	66
2.4	Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Technology-Improvement Sensitivity Analysis	69
2.5	Expected Annual Operating and Investment Costs In Business-As-Usual Case and With Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Technology-Improvement Sensitivity Analysis	71
2.6	Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Transmission-Congestion Sensitivity Analysis	72
2.7	Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Land-Use Sensitivity Analysis .	74
2.8	Capacity built in business-as-usual case and with different carbon-emissions reductions relative to business-as-usual case under PV-solar-and-BES sensitivity analysis	75

LIST OF FIGURES

2.9	Capacity Built In Business-As-Usual Case and with Different Carbon-Emissions Reductions Relative to Business-As-Usual Case Under Adiabatic-CAES Sensitivity Analysis	77
2.10	Capacity built in business-as-usual case and with different carbon-emissions reductions relative to business-as-usual case under electricity-demand sensitivity analysis.	78
2.11	Expected annual operating and investment costs in business-as-usual case and with different carbon-emissions reductions relative to business-as-usual case under electricity-demand sensitivity analysis.	79

Chapter 3

3.1	Iterative solution approach of the ARO problem	114
3.2	69-bus system	121
3.3	Annual ton CO ₂ emissions in the considered cases	133
3.4	69-bus system under the “8CN” the network configuration assumption	135

Chapter 4

4.1	Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Gas NBP time series	158
4.2	ACF of the squared mean adjusted log return series and ACF of the squared mean adjusted residuals of Gas NBP log return series.	159
4.3	ACF and PACF of Gas NBP log return series	160
4.4	ACF and PACF of Gas NBP residuals	161
4.5	First tree of the C -vine with a 5% confidence level. The letters reported between the root nodes indicate the type of the bivariate copulas used to model the dependence, while the numbers refer to the corresponding Kendall’s τ correlation.	163
4.6	First tree of the D -vine with a 5% confidence level. The letters reported between the root nodes indicate the type of the bivariate copulas used to model the dependence, while the numbers refer to the corresponding Kendall’s τ correlation.	164
4.7	First tree of the R -vine with a 5% confidence level. The letters reported between the root nodes indicate the type of the bivariate copulas used to model the dependence, while the numbers refer to the corresponding Kendall’s τ correlation.	165

Chapter 4

4.A.1 Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Brent series	175
4.A.2 Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Gasoil series.	176
4.A.3 Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the JetF series.	177
4.A.4 Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Naphtha series.	178
4.A.5 Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Lsfo series.	179
4.A.6 Log returns (a), 30-days horizon rolling standard deviation on log returns (b), and volatility (c) associated with the Gas HenryHub series.	180
4.A.7 ACF of the squared mean adjusted log return series and ACF of the squared mean adjusted residuals of Brent (a), Gasoil (b), and JetF (c) log return series.	181
4.A.8 ACF of the squared mean adjusted log return series and ACF of the squared mean adjusted residuals of Naphtha (a), Lsfo (b), and Gas HenryHub (c) log return series.	182
4.A.9 ACF and PACF of Brent (a), Gasoil (b), and JetF (c) log return series	183
4.A.10 ACF and PACF of Naphtha (a), Lsfo (b), and Gas HenryHub (c) log return series.	184
4.A.11 ACF and PACF of Brent (a), Gasoil (b), and JetF (c) residuals.	185
4.A.12 ACF and PACF of Naphtha (a), Lsfo (b), and Gas HenryHub (c) residuals.	186

Chapter 5

5.1 Natural gas supply chain	191
5.2 Yearly volume of gas exchanged per type and yearly demand per consumer group (mcm/year)	218
5.3 Natural gas and LNG bought through LTCs (mcm/day)	220
5.4 Natural gas and LNG bought on the spot market (mcm/day)	220
5.5 Yearly volume of gas exchanged per type (mcm/year)	224
5.6 Volumes of gas exchanged via LTCs under the “Risk” and “No FLEX” assumptions (mcm/day)	224
5.7 Volumes of LNG exchanged via LTCs under the “Risk” and “No FLEX” assumptions (mcm/day)	225
5.8 Volumes of spot gas exchanged under the “Risk” and “No FLEX” assumptions (mcm/day)	225

LIST OF FIGURES

5.9	Volumes of natural gas exchanged via LTCs under the “Risk” and “No FLEX” assumptions (mcm/day)	226
5.10	Volumes of natural gas exchanged via LTCs under the “Risk” and “FLEX” assumptions (mcm/day)	227
5.11	Volumes of spot gas exchanged under the “Risk” and “FLEX” assumptions (mcm/day)	228
5.12	Volumes of spot LNG exchanged under the “Risk” and “FLEX” assumptions (mcm/day)	228
5.13	Yearly demand per consumer group (mcm/year)	229

List of Tables

Chapter 2

2.1 Data For Candidate Thermal Units	57
2.2 Social cost of carbon emission	66

Chapter 2

2.A.1 Technical Characteristics of Onshore-Wind Units	83
2.A.2 Technical Characteristics of Offshore-Wind Units	84
2.A.3 Technical Characteristics of CSP Units	84
2.A.4 Technical Characteristics of Diabatic-CAES Units	84
2.A.5 Technical Characteristics of Existing Transmission Lines	85
2.A.6 Technical Characteristics of Candidate ac Transmission Lines	86
2.A.7 Technical Characteristics of Candidate HVDC Transmission Lines	86

Chapter 2

2.B.1 Investment Costs of Candidate HVDC Transmission Lines	87
2.B.2 Technical Characteristics of Additional CSP Units in Land-Use Sensitivity Case . . .	88
2.B.3 Technical Characteristics of Additional CSP Units in Land-Use Sensitivity Case With Technology Improvements	89
2.B.4 Technical Characteristics of Adiabatic-CAES Units With High Costs	89
2.B.5 Technical Characteristics of Adiabatic-CAES Units With Low Costs	90
2.B.6 Technical Characteristics of BES Units	91
2.B.7 Technical Characteristics of PV Units	91

Chapter 3

3.1 Data for candidate units.	120
---------------------------------------	-----

LIST OF TABLES

3.2	Investment decisions for the deterministic case	125
3.3	Impact of the long-term uncertainty of the value of the electricity from the main grid on investment decisions	125
3.4	Impact of the long-term uncertainty of peak demand on investment decisions . . .	127
3.5	Impact of the long-term uncertainty of the number of EVs on investment decisions	127
3.6	Execution time of the long-term uncertainty analysis with the ARO approach . . .	128
3.7	Impact of the revenues associated with EVs on investment decisions	128
3.8	Impact of the limits of the power purchased from main grid on investment decisions	130
3.9	Impact of the budget availability on investment decisions	131
3.10	Results for the deterministic case	136
3.11	Impact of the long-term uncertainty of peak demand on investment decisions . . .	136
3.12	Impact of the long-term uncertainty of the number of EVs on investment decisions	137
3.13	CPU times of the stochastic programming model depending on the number of scenarios	141

Chapter 3

3.A.1	69-buses network data	146
3.A.2	Average peak load values per node	147

Chapter 4

4.1	Basic statistics of log return time series referred to the period January 4, 2012 - July 24, 2014	157
4.2	Unit root test results for the considered series	157
4.3	p-values of ARMA GARCH models. The corresponding statistics are reported in parentheses. WLB=Weighted Ljung-Box Test, WALM= Weighted Arch LM Test. S.R= Standardized Residuals. S.S.R= Standardized Squared Residuals	162
4.4	Kendall's τ correlation between u-data	162
4.5	Log-likelihood, AIC and BIC for the estimated copula models	166
4.6	Goodness-of-fit test on C -vine, D -vine and R -vine with bootstrap repetition rate $x = 200$	166
4.7	Vuong test results at level $\alpha = 5\%$	166
4.8	Clark test results	167
4.9	Optimal weights for long-term natural gas portfolio C -vine	171

4.10	Optimal weights for long-term natural gas portfolio <i>D</i> -vine	172
4.11	Optimal weights for long-term natural gas portfolio <i>R</i> -vine	172
4.12	Oil and gas composition of the optimal portfolios	173

Chapter 5

5.1	Summary of the new indicators proposed	200
5.2	External supply risk indicators	216
5.3	Volumes of gas and LNG regulated by LTCs between Italy and supplying countries in 2015 (annual value)	217
5.4	Supplying countries and mid-streamer's profits, final consumers' surplus, and social welfare under the "NO Risk" assumption (€/year)	221
5.5	Prices of gas and LNG LTCs under the "NO Risk" assumption (€/cm)	222
5.6	Prices spot of gas and LNG under the "NO Risk" assumption (€/cm)	222
5.7	Final consumers' prices under the "NO Risk" assumption (€/cm)	223
5.8	Supplying countries that exchange gas and LNG with Italy under the "Risk" and "FLEX" assumptions	232
5.9	Supplying countries and mid-streamer's profits, final consumers' surplus, and social welfare under the "NO Risk" assumption (€/year)	233
5.10	Prices of gas LTCs under the "Risk FLEX" assumptions (€/cm)	233
5.11	Final consumers' prices under the "Risk FLEX" assumptions (€/cm)	233

Bibliography

- Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44, 182–198.
- Abada, I., A. Ehrenmann, and Y. Smeers (2017). Modeling gas markets with endogenous long-term contracts. *Operations research* 65(4), 856–877.
- Accioly, R. M. S. and F. A. L. Aiube (2008). Analysis of crude oil and gasoline prices through copulas. *Cadernos do IME- Série Estatística* 24, 5–28.
- Ackermann, T., G. Andersson, and L. Söder (2001). Distributed generation: a definition1. *Electric power systems research* 57(3), 195–204.
- Agency for the Cooperation of Energy Regulators (2015). *European Gas Target Model review and update*. Trg republike 3, 1000 Ljubljana, Slovenia: Agency for the Cooperation of Energy Regulators.
- Aloui, R., M. S. B. Aïssa, and D. K. Nguyen (2013). Conditional dependence structure between oil prices and exchange rates: a copula-garch approach. *Journal of International Money and Finance* 32, 719–738.
- Amjady, N., A. Attarha, S. Dehghan, and A. J. Conejo (2017). Adaptive robust expansion planning for a distribution network with ders. *IEEE Transactions on Power Systems* 33(2), 1698–1715.
- Asensio, M., P. M. de Quevedo, G. Muñoz-Delgado, and J. Contreras (2016). Joint distribution network and renewable energy expansion planning considering demand response and energy storage—part i: Stochastic programming model. *IEEE Transactions on Smart Grid* 9(2), 655–666.

- Autorità per l'Energia Elettrica il Gas e il Sistema Idrico (2016). *Relazione annuale sullo stato dei servizi e sull'attività svolta*. Corso di Porta Vittoria, 27 - 20122 Milano: Autorità per l'Energia Elettrica il Gas e il Sistema Idrico.
- Baran, M. E. and F. F. Wu (1989). Optimal capacitor placement on radial distribution systems. *IEEE Transactions on power Delivery* 4(1), 725–734.
- Barbour, E., D. Mignard, Y. Ding, and Y. Li (2015, 1 October). Adiabatic Compressed Air Energy Storage with packed bed thermal energy storage. *Applied Energy* 155, 804–815.
- Baringo, A. and L. Baringo (2016). A stochastic adaptive robust optimization approach for the offering strategy of a virtual power plant. *IEEE Transactions on Power Systems* 32(5), 3492–3504.
- Baringo, L. and A. Baringo (2017). A stochastic adaptive robust optimization approach for the generation and transmission expansion planning. *IEEE Transactions on Power Systems* 33(1), 792–802.
- Baringo, L. and A. Conejo (2013a). Correlated wind-power production and electric load scenarios for investment decisions. *Applied energy* 101, 475–482.
- Baringo, L. and A. J. Conejo (2010). Offering strategy via robust optimization. *IEEE Transactions on Power Systems* 26(3), 1418–1425.
- Baringo, L. and A. J. Conejo (2013b, January). Correlated wind-power production and electric load scenarios for investment decisions. *Applied Energy* 101, 475–482.
- Barnes, F. S., D. A. Budd, M. Lim, and E. R. Freeman (2015, June). Compressed Air Energy Storage (CAES). In L. F. Cabeza, R. Sioshansi, and J. Yan (Eds.), *Handbook of Clean Energy Systems*, Volume 5, Energy Storage, Chapter 19, pp. 2717–2742. West Sussex, United Kingdom: John Wiley & Sons Ltd.
- Bassetti, F., M. E. De Giuli, E. Nicolino, and C. Tarantola (2018). Multivariate dependence analysis via tree copula models: An application to one-year forward energy contracts. *European Journal of Operational Research*, 1–15.
- Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32, 245–268.

BIBLIOGRAPHY

- Bedford, T. and R. M. Cooke (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, 1031–1068.
- Bekiros, S., J. A. Hernandez, S. Hammoudeh, and D. K. Nguyen (2015). Multivariate dependence risk and portfolio optimization: An application to mining stock portfolios. *Resources Policy* 46, 1–11.
- Ben-Tal, A., L. El Ghaoui, and A. Nemirovski (2009). *Robust optimization*, Volume 28. Princeton University Press.
- Bertsimas, D. and D. B. Brown (2009). Constructing uncertainty sets for robust linear optimization. *Operations Research*, 57(6), 1483–1495.
- Bertsimas, D., D. B. Brown, and C. Caramanis (2011). Theory and applications of robust optimization. *SIAM Review*, 53(3), 464–501.
- Bertsimas, D., E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng (2012). Adaptive robust optimization for the security constrained unit commitment problem. *IEEE transactions on power systems* 28(1), 52–63.
- Bertsimas, D. and M. Sim (2004). The price of robustness. *Operations Research*, 52, 35–53.
- Bilgili, M., A. Yasar, and E. Simsek (2011). Offshore wind power development in europe and its comparison with onshore counterpart. *Renewable and Sustainable Energy Reviews* 15(2), 905–915.
- Birge, J. R. and F. Louveaux (1997). *Introduction to Stochastic Programming*. New York, NY: Springer-Verlag.
- Blair, N., A. P. Dobos, J. Freeman, T. Neises, M. Wagner, T. Ferguson, P. Gilman, and S. Janzou (2014, February). System Advisor Model, SAM 2014.1.14: General Description. Technical Report NREL/TP-6A20-61019, National Renewable Energy Laboratory.
- Blyth, W. and N. Lefevre (2004). *Energy Security and Climate Change: An Assessment Framework*. International Energy Agency & Organisation for Economic Co-operation and Development.
- BP (2016). *Statistical Review of World Energy 2016-Natural gas*. 1 St James’s Square, London, United Kingdom.

- Brechmann, E. C. and C. Czado (2013). Risk management with high-dimensional vine copulas: an analysis of the euro stoxx 50. *Stat. Risk Model* 30, 307–342.
- Calderón, S., A. C. Alvarez, A. M. Loboguerrero, S. Arango, K. Calvin, T. Kober, K. Daenzer, and K. Fisher-Vanden (2016, May). Achieving CO₂ reductions in Colombia: Effects of carbon taxes and abatement targets. *Energy Economics* 56, 575–586.
- Chekhlov, A., S. Uryasev, and M. Zabarankin (2005). Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance* 8, 13–58.
- Chyong, C. K. (2015). Markets and long-term contracts: The case of russian gas supplies to europe. Working Paper 1524, EPRG.
- Clarke, K. A. (2007). A simple distribution-free test for nonnested model selection. *Political Analysis* 15, 347–363.
- Cohen, G., F. Joutz, and P. Loungani (2011). Measuring energy security: Trends in the diversification of oil and natural gas supplies. *Energy policy* 39(9), 4860–4869.
- Conejo, A. J., M. Carrión, and J. M. Morales (2010). *Decision making under uncertainty in electricity markets*, Volume 1. Springer.
- Conejo, A. J., L. B. Morales, S. J. Kazempour, and A. S. Siddiqui (2016). *Investment in Electricity Generation and Transmission: Decision Making under Uncertainty*, Volume 205. Cham, Switzerland: Springer International Publishing.
- Contreras, J., M. Asensio, P. M. de Quevedo, G. Muñoz-Delgado, and S. Montoya-Bueno (2016). *Joint RES and Distribution Network Expansion Planning under a Demand Response Framework*. Academic Press.
- Cooper, A. and K. Schefter (2017, July). Plug-in Electric Vehicle Sales Forecast Through 2025 and the Charging Infrastructure Required. Technical report, Institute for Electric Innovation and Edison Electric Institute.
- Correljé, A. (2016). The european natural gas market. *Current Sustainable/Renewable Energy Reports* 3(1-2), 28–34.

BIBLIOGRAPHY

- Czado, C., A. Min, T. Baumann, and R. Dakovic (2009). Pair-copula constructions for modeling exchange rate dependence. Preprint.
- Czado, C., U. Schepsmeier, and A. Min (2012). Maximum likelihood estimation of mixed c-vines with application to exchange rates. *Statistical Modelling* 12(3), 229–255.
- Dalla Valle, L., M. E. D. Giuli, C. Tarantola, and C. Manelli (2016). Default probability estimation via pair-copula constructions. *European Journal of Operational Research* 249, 298–311.
- Das, T., V. Krishnan, Y. Gu, and J. D. McCalley (2011, 24-29 July). Compressed Air Energy Storage: State Space Modeling and Performance Analysis. In *2011 Power and Energy Society General Meeting*, San Diego, CA. Institute of Electrical and Electronics Engineers.
- de Maere d’Aertrycke, G., E. Ehrenmann, and Y. Smeers (2017). Investment with incomplete markets for risk: The need for long-term contracts. *Energy Policy* 105, 571–583.
- de Quevedo, P. M., G. Munoz-Delgado, and J. Contreras (2017). Impact of electric vehicles on the expansion planning of distribution systems considering renewable energy, storage, and charging stations. *IEEE Transactions on Smart Grid* 10(1), 794–804.
- Denholm, P., E. Ela, B. Kirby, and M. R. Milligan (2010, January). The Role of Energy Storage with Renewable Electricity Generation. Technical Report NREL/TP-6A2-47187, National Renewable Energy Laboratory.
- Di Sbroiavacca, N., G. Nadal, F. Lallana, J. Falzon, and K. Calvin (2016, May). Emissions reduction scenarios in the Argentinean Energy Sector. *Energy Economics* 56, 552–563.
- Dickey, D. and W. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49, 1057–1072.
- Dirkse, S. P. and M. C. Ferris (1995). The path solver: a nonmonotone stabilization scheme for mixed complementarity problems. *Optimization Methods and Software* 5(2), 123–156.
- Dißmann, J., E. Brechmann, C. Czado, and D. Kurowicka (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis* 59, 52–69.

- Divya, K. and J. Østergaard (2009). Battery energy storage technology for power systems—an overview. *Electric Power Systems Research* 79(4), 511–520.
- Dobrota, D., B. Lalić, and I. Komar (2013). Problem of boil-off in lng supply chain. *Transactions on maritime science* 2(02), 91–100.
- Dominguez, R., L. Baringo, and A. Conejo (2012). Optimal offering strategy for a concentrating solar power plant. *Applied Energy* 98, 316–325.
- Domínguez, R., A. J. Conejo, and M. Carrión (2015, January). Toward Fully Renewable Electric Energy Systems. *IEEE Transactions on Power Systems* 30, 316–326.
- Draxl, C., B.-M. Hodge, A. Clifton, and J. McCaa (2015, April). Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit. Technical Report NREL/TP-5000-61740, National Renewable Energy Laboratory.
- Drury, E., P. Denholm, and R. Sioshansi (2011, August). The Value of Compressed Air Energy Storage in Energy and Reserve Markets. *Energy* 36, 4959–4973.
- Durante, F. and C. Sempì (2015). *Principles of copula theory*. Chapman and Hall/CRC: London.
- Ebrahimi, M. R. and N. Amjady (2019). Adaptive robust optimization framework for day-ahead microgrid scheduling. *International Journal of Electrical Power & Energy Systems* 107, 213–223.
- Egging, R. (2013). Benders decomposition for multi-stage stochastic mixed complementarity problems—applied to a global natural gas market model. *European Journal of Operational Research* 226(2), 341–353.
- Egging, R., S. A. Gabriel, F. Holz, and J. Zhuang (2008). A complementarity model for the european natural gas market. *Energy policy* 36(7), 2385–2414.
- Egging, R. and F. Holz (2016). Risks in global natural gas markets: investment, hedging and trade. *Energy Policy* 94, 468–479.
- Egging, R., F. Holz, and S. A. Gabriel (2010). The world gas model: A multi-period mixed complementarity model for the global natural gas market. *Energy* 35(10), 4016–4029.
- Egging, R. G. and S. A. Gabriel (2006). Examining market power in the european natural gas market. *Energy Policy* 34(17), 2762–2778.

BIBLIOGRAPHY

- El Khatib, S. and F. D. Galiana (2007). Negotiating bilateral contracts in electricity markets. *IEEE Transactions on Power Systems* 22(2), 553–562.
- European Commission (2011a). *Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee, the Committee of the Regions and the Committee of the Regions. Energy Roadmap 2050* (COM(2011) 885 final ed.). European Commission.
- European Commission (2011b). *White Paper, Roadmap to a Single European Transport Area-Towards a competitive and resource efficient transport system* (COM(2011) 144 final ed.). European Commission.
- European Commission (2016). *Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee, the Committee of the Regions and the Committee of the Regions. A European Strategy for Low-Emission Mobility*. (COM(2016) 501 final ed.). European Commission.
- European Network of Transmission System Operator for Gas (2012). *Balancing tools-product definition and deployment*. Avenue de Cortenbergh 100 / Kortenberglaan 100, Bruxelles, Belgium: European Network of Transmission System Operator for Gas.
- Facchinei, F. and J.-S. Pang (2007). *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media.
- Feldman, D. J. and R. M. Margolis (2018). Q2/q3 2018 solar industry update. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Fischer, R., P. Serra, P. L. Joskow, and W. W. Hogan (2000). Regulating the electricity sector in latin america [with comments]. *Economia* 1(1), 155–218.
- Fisher, T. J. and C. M. O. Gallagher (2012). New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Association* 107(498), 777–787.
- Flavin, C. and S. Kitasei (2010). *The role of natural gas in a low-carbon energy economy*. Worldwatch Institute Washington, DC.
- Franza, L. (2014). Long-term gas import contracts in europe. *CIEP paper* 8.

- Frisch, M. (2010). Current european gas pricing problems: solutions based on price review and price reopener provisions. Research paper Series, No 3, International energy law and policy.
- Fronzel, M. and C. M. Schmidt (2008). Measuring energy security-a conceptual note. *Ruhr Economic Paper No. 52*.
- Fu, R., D. Feldman, R. Margolis, M. Woodhouse, and K. Ardani (2017). Us solar photovoltaic system cost benchmark: Q1 2017. Technical report, EERE Publication and Product Library.
- Fu, R., T. W. Remo, and R. M. Margolis (2018). 2018 US Utility-Scale Photovoltaics-Plus-Energy Storage System Costs Benchmark. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Gabriel, S. A., A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz (2013). Optimality and complementarity. In *Complementarity Modeling in Energy Markets*, pp. 31–69. Springer.
- García-Bertrand, R. and R. Mínguez (2016). Dynamic robust transmission expansion planning. *IEEE Transactions on Power Systems* 32(4), 2618–2628.
- García-Cerezo, Á., L. Baringo, and R. García-Bertrand (2020). Representative days for expansion decisions in power systems. *Energies* 13(2), 335.
- Gaupp, F., G. Pflug, S. Hochrainer-Stigler, J. Hall, and S. Dadson (2017). Dependency of crop production between global breadbaskets: a copula approach for the assessment of global and regional risk pools. *Risk Analysis* 37(11), 2212–2228.
- Genest, C. and A. C. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12, 347–368.
- Genest, C. M. G. and M. Bourdeau-Brien (2009). The advent of copulas in finance. *The European Journal of Finance* 15, 609–618.
- Gestore Mercati Energetici (2009). *Vademecum della borsa elettrica*. Viale Maresciallo Pilsudski, 122/124 Roma, Italia: Gestore Mercati Energetici.
- Gibbins, J. and H. Chalmers (2008). Carbon capture and storage. *Energy policy* 36(12), 4317–4322.
- Global Wind Energy Council (2016). *Global wind energy outlook 2016*. 80 Rue d’Arlon, 1040 Brussels, Belgium: Global Wind Energy Council.

BIBLIOGRAPHY

- Global Wind Energy Council (2017). *Global Wind Report: Annual market up date 2017*. 80 Rue d'Arlon, 1040 Brussels, Belgium: Global Wind Energy Council.
- Golestaneh, F., P. Pinson, R. Azizipanah-Abarghooee, and H. B. Gooi (2018). Ellipsoidal prediction regions for multivariate uncertainty characterization. *IEEE Transactions on Power Systems* 33(4), 4519–4530.
- Golestaneh, F., P. Pinson, and H. B. Gooi (2018). Polyhedral predictive regions for power system applications. *IEEE Transactions on Power Systems* 34(1), 693–704.
- Gomez-Exposito, A., A. J. Conejo, and C. Canizares (2018). *Electric energy systems: analysis and operation*. CRC press.
- Graf, C. and C. Marcantonini (2017, August). Renewable energy and its impact on thermal generation. *Energy Economics* 66, 421–430.
- Graves, F., T. Jenkin, and D. Murphy (1999, October). Opportunities for Electricity Storage in Deregulating Markets. *The Electricity Journal* 12, 46–56.
- Greenblatt, J. B., S. Succar, D. C. Denkenberger, R. H. Williams, and R. H. Socolow (2007, March). Baseload wind energy: modeling the competition between gas turbines and compressed air energy storage for supplemental generation. *Energy Policy* 35, 1474–1492.
- Grégoire, V., C. Genest, and M. Gendron (2008). Using copulas to model price dependencies in energy markets. *Energy Risk* 5, 58–64.
- Grubb, M., L. Butler, and P. Twomey (2006). Diversity and security in uk electricity generation: The influence of low-carbon objectives. *Energy policy* 34(18), 4050–4062.
- Gupta, E. (2008). Oil vulnerability index of oil-importing countries. *Energy policy* 36(3), 1195–1211.
- Hartmann, N., O. Vöhringer, C. Kruck, and L. Eltrop (2012, May). Simulation and analysis of different adiabatic Compressed Air Energy Storage plant configurations. *Applied Energy* 93, 541–548.
- Hasan, M. F., A. M. Zheng, and I. Karimi (2009). Minimizing boil-off losses in liquefied natural gas transportation. *Industrial & engineering chemistry research* 48(21), 9571–9580.
- Heather, P. (2012). Continental european gas hubs: Are they fit for purpose? Technical Report 63, NG: Oxford Institute for Energy Studies.

- Heather, P. (2015). The evolution of european traded gas hubs. Technical report, Oxford Institute for Energy Studies Oxford.
- Heather, P. and B. Petrovich (2017). European traded gas hubs: an updated analysis on liquidity, maturity and barriers to market integration. Technical report, The Oxford Institute for Energy Studies paper-Energy Insight.
- Hernandez, J. A. (2014). Are oil and gas stocks from the australian market riskier than coal and uranium stocks? dependence risk analysis and portfolio optimization. *Energy Economics* 45, 528–536.
- Herrero, I., P. Rodilla, C. Batlle, et al. (2018). Enhancing intraday price signals in us iso markets for a better integration of variable energy resources. *The Energy Journal* 39(3).
- Hofert, M., I. Kojadinovic, M. Maechler, and J. Yan (2017). *Copula*. R package.
- Holz, F., H. Engerer, C. Kemfert, P. M. Richter, and C. von Hirschhausen (2014). *European natural gas infrastructure: The role of Gazprom in European natural gas supplies. Study Commissioned by The Greens/European Free Alliance in the European Parliament*. Number 81. DIW Berlin: Politikberatung kompakt.
- Holz, F., P. Richter, and R. Egging-Bratseth (2013). The role of natural gas in a low-carbon europe: Infrastructure and regional supply security in the global gas model. Discussion Papers of DIW Berlin 1273, DIW Berlin, German Institute for Economic Research.
- Holz, F., P. M. Richter, and R. Egging (2016). The role of natural gas in a low-carbon europe: Infrastructure and supply security. *Energy Journal* 37.
- Honoré, A. (2013). The italian gas market: Challenges and opportunities. Technical Report 76, Oxford Institute for Energy Studies Oxford.
- Hull, J. C. and S. Basu (2016). *Options, futures, and other derivatives*. Pearson Education India.
- International Energy Agency (2007). *Energy Security and Climate Change; assessing interactions*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2013). *Developing a Natural Gas Trading Hub in Asia-Obstacles and Opportunities*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.

BIBLIOGRAPHY

- International Energy Agency (2014a). *Techology Roadmap: Solar Photovoltaic Energy*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2014b). *Techology Roadmap: Solar Thermal Electricity*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2016a). *Annual Gas Statistic Database*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2016b, 25 October). *Medium-Term Renewable Energy Market Report 2016*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2016c). *Updated Capital Cost Estimates for Utility Scale Electricity Generating Plants*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2016d). *World Energy Outlook 2016*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency (2018). *Annual Energy Outlook 2018*. 9 rue de la Fédération, 75739 Paris Cedex 15, France: International Energy Agency.
- International Energy Agency & Organisation for Economic Co-operation and Development (2001). *Toward a sustainable energy future*. Paris: International Energy Agency & Organisation for Economic Co-operation and Development.
- International Group of Liquefied Natural Gas Importers (2018). *The LNG industry*. International Group of Liquefied Natural Gas Importers.
- International Renewable Energy Agency (2018a). *Global Energy Transformation: A Roadmap to 2050*. Masdar City, P.O. Box 236, Abu Dhabi, United Arab Emirates: International Renewable Energy Agency.
- International Renewable Energy Agency (2018b). *Renewable Power Generation Costs in 2017*. Masdar City, P.O. Box 236, Abu Dhabi, United Arab Emirates: International Renewable Energy Agency.
- Jabr, R. A. (2013). Robust transmission network expansion planning with uncertain renewable generation and loads. *IEEE Transactions on Power Systems*, 28(4), 4558–4567.

- Jäschke, S. (2014). Estimation of risk measures in energy portfolios using modern copula techniques. *Computational Statistics and Data Analysis* 76, 359–376.
- Jess, M. (1997). Restructuring energy industries: lessons from natural gas. *Natural Gas Monthly Special Report*.
- Jiang, R., J. Wang, and Y. Guan (2011). Robust unit commitment with wind power and pumped storage hydro. *IEEE Transactions on Power Systems* 27(2), 800–810.
- Jiang, R., M. Zhang, G. Li, and Y. Guan (2014). Two-stage network constrained robust unit commitment problem. *European Journal of Operational Research* 234(3), 751–762.
- Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. *IMS lecture notes* 76, 359–376.
- Joe, H. (1997). “Multivariate model and dependence concepts”. 73, Chapman, Hall, London: Monographs on Statistics and Applied Probability.
- Jónsson, T., P. Pinson, and H. Madsen (2010). On the market impact of wind energy forecasts. *Energy Economics* 32(2), 313–320.
- Kall, P., S. W. Wallace, and P. Kall (1994). *Stochastic programming*. Springer.
- Karan, M. B. and H. Kazdağlı (2011). The development of energy markets in europe. In *Financial Aspects in Energy*, pp. 11–32. Springer.
- Kazempour, S. J., A. J. Conejo, and C. Ruiz (2013). Generation investment equilibria with strategic producers—part i: Formulation. *IEEE Transactions on Power Systems* 28(3), 2613–2622.
- Khaligh, V. and A. Anvari-Moghaddam (2019). Stochastic expansion planning of gas and electricity networks: A decentralized-based approach. *Energy* 186, 115889.
- Khaligh, V., M. Oloomi Buygi, A. Anvari-Moghaddam, and J. M Guerrero (2018). A multi-attribute expansion planning model for integrated gas–electricity system. *Energies* 11(10), 2573.
- King, J., A. Clifton, and B.-M. Hodge (2014, September). Validation of Power Output for the WIND Toolkit. Technical Report NREL/TP-5D00-61714, National Renewable Energy Laboratory.

BIBLIOGRAPHY

- Konno, H., H. Shirakawa, and H. Yamazaki (1993). A mean-absolute deviation-skewness portfolio optimization models. *Annals of Operations Research* 45(1), 205–220.
- Krohn, S., P.-E. Morthorst, and S. Awerbuch (2009, March). The Economics of Wind Energy. Technical report, European Wind Energy Association.
- Kruyt, B., D. P. van Vuuren, H. J. de Vries, and H. Groenenberg (2009). Indicators for energy security. *Energy policy* 37(6), 2166–2181.
- Krzemienowski, A. and S. Szymczyk (2016). Portfolio optimization with a copula-based extension of conditional value-at-risk. *Annals of Operations Research* 237(1-2), 219–236.
- Kurowicka, D. and R. Cooke (2006). “*Uncertainty analysis with high dimensional dependence modelling*”. Wiley: Chichester.
- Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unitroot. *Journal of Econometrics* 54, 159–178.
- Laing, T., M. Sato, M. Grubb, C. Combetti, et al. (2013). Assessing the effectiveness of the eu emissions trading system. *Centre for Climate Change Economics and Policy Working Paper* 126.
- Le Coq, C. and E. Paltseva (2009). Measuring the security of external energy supply in the european union. *Energy Policy* 37(11), 4474–4481.
- Lee, W.-W. (2004). Us lessons for energy industry restructuring: based on natural gas and california electricity incidences. *Energy Policy* 32(2), 237–259.
- Li, X. and Y. You (2014). A note on allocation of portfolio shares of random assets with Archimedean copula”. *Annals of Operations Research* 212(1), 155–167.
- Lilliestam, J., M. Labordena, A. Patt, and S. Pfenninger (2017, July). Empirically observed learning rates for concentrating solar power and their responses to regime change. *Nature Energy* 2, 1–6.
- Lima, R. M., A. Q. Novais, and A. J. Conejo (2015). Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. an adaptive robust optimization approach. *European Journal of Operational Research* 240(2), 457–475.
- Liu, J.-L. and J.-H. Wang (2016, 15 January). A comparative research of two adiabatic compressed air energy storage systems. *Energy Conversion and Management* 108, 566–578.

- Liu, Y., R. Sioshansi, and A. J. Conejo (2018a, May). Hierarchical Clustering to Find Representative Operating Periods for Capacity-Expansion Modeling. *IEEE Transactions on Power Systems* 33, 3029–3039.
- Liu, Y., R. Sioshansi, and A. J. Conejo (2018b, January). Multistage Stochastic Investment Planning with Multiscale Representation of Uncertainties and Decisions. *IEEE Transactions on Power Systems* 33, 781–791.
- Lu, X. F., K. K. Lai, and L. Liang (2014). Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model. *Annals of Operations Research* 219, 333–357.
- Lumbreras, S. and A. Ramos (2016). The new challenges to transmission expansion planning. survey of recent practice and literature review. *Electric Power Systems Research* 134, 19–29.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, pp. 281–297.
- Madaeni, S. H., R. Sioshansi, and P. Denholm (2012, February). How Thermal Energy Storage Enhances the Economic Viability of Concentrating Solar Power. *Proceedings of the IEEE* 100, 335–347.
- Mai, T., J. Paige, J. Logan, C. McMillan, M. Muratori, D. Steinberg, L. Vimmerstedt, R. Jones, B. Haley, and N. Brent (2018). Electrification Futures Study: Scenarios of Electric Technology Adoption and Power Consumption for the United States. Technical Report NREL/TP-6A20-71500, National Renewable Energy Laboratory.
- Mandelbrot, B. B. (1963). The variation of certain speculative prices. *The Journal of Business* 36, 394–419.
- Markowitz, H. (1952). Portfolio selection. *The journal of finance* 7(1), 77–91.
- Mauritsen, T. and R. Pincus (2017). Committed warming inferred from observations. *Nature Climate Change*.
- Melling, A. J. (2010). Natural gas pricing and its future-europe ad the battleground. Technical report, Carnegie Endowment for international peace.

BIBLIOGRAPHY

- Mínguez, R. and R. García-Bertrand (2016). Robust transmission network expansion planning in energy systems: Improving computational performance. *European Journal of Operational Research*, 248 (1), 21–32.
- Morales, J. M., A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno (2013). *Integrating renewables in electricity markets: operational problems*, Volume 205. Springer Science & Business Media.
- Morales-Napoles, O. (2010). Counting vines. *Dependence modeling: vine copula handbook*, 189–218.
- Moreira, A., A. Street, and J. M. Arroyo (2015). Energy and reserve scheduling under correlated nodal demand uncertainty: An adjustable robust optimization approach. *International Journal of Electrical Power & Energy Systems* 72, 91–98.
- Moreno, R., A. Street, J. M. Arroyo, and P. Mancarella (2017). Planning low-carbon electricity systems under uncertainty considering operational flexibility and smart grid technologies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375(2100), 20160305.
- Munoz-Delgado, G., J. Contreras, and J. M. Arroyo (2019). Distribution system expansion planning considering non-utility-owned dg and an independent distribution system operator. *IEEE Transactions on Power Systems* 34(4), 2588–2597.
- Nagurney, A. (1999). *Network economics: A variational inequality approach* (revised second edition) kluwer academic publishers.
- Nair, N.-K. C. and N. Garimella (2010). Battery energy storage systems: Assessment for small-scale renewable energy integration. *Energy and Buildings* 42(11), 2124–2130.
- Najafi, J., A. Peiravi, A. Anvari-Moghaddam, and J. M. Guerrero (2019). Resilience improvement planning of power-water distribution systems with multiple microgrids against hurricanes using clean strategies. *Journal of cleaner production* 223, 109–126.
- Navidi, M., S. M. Moghaddas Tafreshi, and A. Anvari-Moghaddam (2019). Sub-transmission network expansion planning considering regional energy systems: A bi-level approach. *Electronics* 8(12), 1416.
- Neises, T. W. and C. S. Turchi (2014). A Comparison of Supercritical Carbon Dioxide Power Cycle Configurations with an Emphasis on CSP Applications. *Energy Procedia* 49, 1187–1196.

- Nelsen, R. B. (1999). *An introduction to copulas*. New York: Springer-Verlag Lecture Notes in Statistics, 139.
- NERA economic consulting (2014). *Updated Macroeconomic Impacts of LNG Exports from the United States*. NERA economic consulting.
- Neumann, A. (2004). Security of supply in liberalised european gas markets. *Europa-University Viadrina Frankfurt/Oder*, 17–20.
- Pandžić, H., Y. Wang, T. Qiu, Y. Dvorkin, and D. S. Kirschen (2015). Near-optimal method for siting and sizing of distributed storage in a transmission network. *IEEE Transactions on Power Systems*, 30(5), 2288–2300.
- Peng, C., P. Xie, and C. Chen (2014). Adjustable robust optimal dispatch of power system with large-scale photovoltaic power stations. In *Proceedings of the CSEE*, Volume 34, pp. 4325–4332.
- Phillips, P. C. B. and P. Perron (1988). Testing for unit roots in time series regression. *Biometrika* 75, 335–346.
- Pickard, W. F., A. Q. Shen, and N. J. Hansing (2009). Parking the power: Strategies and physical limitations for bulk energy storage in supply–demand matching on a grid whose input power is provided by intermittent sources. *Renewable and Sustainable Energy Reviews* 13(8), 1934–1945.
- Pineda, S. and J. M. Morales (2018). Chronological time-period clustering for optimal capacity expansion planning with storage. *IEEE Transactions on Power Systems* 33(6), 7162–7170.
- Pineda, S. and J. M. Morales (2019). Solving linear bilevel problems using big-ms: not all that glitters is gold. *IEEE Transactions on Power Systems* 34(3), 2469–2471.
- Prékopa, A. (2013). *Stochastic programming*, Volume 324. Springer Science & Business Media.
- Qi, T., Y. Weng, X. Zhang, and J. He (2016, November). An analysis of the driving factors of energy-related CO₂ emission reduction in China from 2005 to 2013. *Energy Economics* 60, 15–22.
- Quemada, J. M. M., J. García-Verdugo, and G. Escibano (2012). *Energy security for the EU in the 21st century: markets, geopolitics and corridors*. Routledge.

BIBLIOGRAPHY

- Rademaekers, K., A. Slingenberg, and S. Morsy (2008). Review and analysis of eu wholesale energy markets—historical and current data analysis of eu wholesale electricity, gas and co2 markets. *Final Report prepared for the European Commission DG TREN, Rotterdam*.
- Raftery, A. E., A. Zimmer, D. M. Frierson, R. Startz, and P. Liu (2017). Less than 2° C warming by 2100 unlikely. *Nature Climate Change*.
- Reboredo, J. C. (2011). How do crude oil prices co-move?: A copula approach. *Energy Economics* 33, 95–113.
- Rissman, J. (2017, September). The future of electric vehicles in the U.S. Technical report, Energy Innovation.
- Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of risk* 2, 21–42.
- Roldán, C., A. S. de la Nieta, R. García-Bertrand, and R. Mínguez (2018). Robust dynamic transmission and renewable generation expansion planning: walking towards sustainable systems. *International Journal of Electrical Power & Energy Systems* 96, 52–63.
- Ruiz, C. and A. J. Conejo (2015). Robust transmission expansion planning. *European Journal of Operational Research*, 242(2), 390–401.
- Saida, A. B. and J. Prigent (2018). On the robustness of portfolio allocation under copula misspecification. *Annals of Operations Research* 262(2), 631–652.
- Schepsmeier, U. (2016). A goodness-of-fit test for regular vine copula models. *Econometric Reviews*, 1–22.
- Schepsmeier, U. and C. Czado (2016). Dependence modelling with regular vine copula models: a case-study for car crash simulation data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 65(3), 415–429.
- Schepsmeier, U., J. Stoeber, E. C. Brechmann, B. Graeler, T. Nagler, and T. Erhardt (2018). *VineCopula: statistical inference of vine copulas*. R package.
- Sciacovelli, A., Y. Li, H. Chen, Y. Wu, J. Wang, S. Garvey, and Y. Ding (2017, 1 January). Dynamic simulation of Adiabatic Compressed Air Energy Storage (A-CAES) plant with integrated thermal

- storage - Link between components performance and plant performance. *Applied Energy* 185 Part 1, 16–28.
- Sengupta, M., A. Habte, P. Gotseff, A. Weekley, A. Lopez, M. Anderberg, C. Molling, and A. Heidinger (2014, September). A Physics-Based GOES Product for Use in NREL's National Solar Radiation Database. Technical Report NREL/CP-5D00-62776, National Renewable Energy Laboratory.
- Sengupta, M., A. Habte, P. Gotseff, A. Weekley, A. Lopez, C. Molling, and A. Heidinger (2014, July). A Physics-Based GOES Satellite Product for Use in NREL's National Solar Radiation Database. Technical Report NREL/CP-5D00-62237, National Renewable Energy Laboratory.
- Shahidehpour, M., H. Yamin, and Z. Li (2003). *Market operations in electric power systems: forecasting, scheduling, and risk management*. John Wiley & Sons.
- Sijm, J., K. Neuhoﬀ, and Y. Chen (2006). Co2 cost pass-through and windfall profits in the power sector. *Climate policy* 6(1), 49–72.
- Silva, R. A., J. J. West, J.-F. Lamarque, D. T. Shindell, W. J. Collins, G. Faluvegi, G. A. Folberth, L. W. Horowitz, T. Nagashima, V. Naik, et al. (2017). Future global mortality from changes in air pollution attributable to climate change. *Nature Climate Change*.
- Sioshansi, F. P. and W. Pfaffenberger (2006). *Electricity market reform: an international perspective*. Elsevier.
- Sioshansi, R. and P. Denholm (2010a, February). The Value of Concentrating Solar Power and Thermal Energy Storage. Technical Report NREL/TP-6A2-45833, National Renewable Energy Laboratory.
- Sioshansi, R. and P. Denholm (2010b, October). The Value of Concentrating Solar Power and Thermal Energy Storage. *IEEE Transactions on Sustainable Energy* 1, 173–183.
- Sioshansi, R., P. Denholm, and T. Jenkin (2011, January). A Comparative Analysis of the Value of Pure and Hybrid Electricity Storage. *Energy Economics* 33, 56–66.
- Sioshansi, R., P. Denholm, and T. Jenkin (2012, March). Market and Policy Barriers to Deployment of Energy Storage. *Economics of Energy & Environmental Policy* 1, 47–63.

BIBLIOGRAPHY

- Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris* 8, 229–231.
- Smith, M. and J. Castellano (2015). Costs associated with non-residential electric vehicle supply equipment: Factors to consider in the implementation of electric vehicle charging stations. Technical report.
- Snam (2016). *Annual Report*. Piazza Santa Barbara, 7 San Donato Milanese, Milan, Italy.
- Stern, J. and H. V. Rogers (2014). The dynamics of a liberalised european gas market: Key determinants of hub prices, and roles and risks of major players. Technical report, The Oxford Institute for Energy Studies.
- Succar, S., J. B. Greenblatt, D. C. Denkenberger, and R. H. Williams (2006, June 4-7). An Integrated Optimization Of Large-Scale Wind With Variable Rating Coupled To Compressed Air Energy Storage. In *Windpower 2006*, Pittsburgh, Pennsylvania. American Wind Energy Association.
- Succar, S. and R. H. Williams (2008, April). Compressed Air Energy Storage Theory, Resources, And Applications For Wind Power. Technical report, Princeton Environmental Institute.
- Sun, K., M. R. Sarker, and M. A. Ortega-Vazquez (2015). Statistical characterization of electric vehicle charging in different locations of the grid. *2015 IEEE Power & Energy Society General Meeting*.
- Theisen, N. (2014). Natural gas pricing in the eu: From oil-indexation to a hybrid pricing system. Technical report, Regional Centre for Energy Policy Research.
- Tissot, B. and D. Welte (2012). *Petroleum formation and occurrence: a new approach to oil and gas exploration*. Springer Science & Business Media.
- Tran, H. D., U. H. Pham, S. Ly, and T. T. Vo-Duy (2017). Extraction dependence structure of distorted copulas via a measure of dependence. *Annals of Operations Research* 256(2), 221–236.
- Turchi, C. S., Z. Ma, T. W. Neises, and M. J. Wagner (2013, 1 November). Thermodynamic Study of Advanced Supercritical Carbon Dioxide Power Cycles for Concentrating Solar Power Systems. *Journal of Solar Energy Engineering* 135, 041007.
- United Nation Framework Convention on Climate Change (2015). *Paris Agreement*. United Nation Framework Convention on Climate Change.

- U.S. Energy Information Administration (2017, 5 January). *Annual Energy Outlook 2017* (DOE/EIA-0383 (2017) ed.). U.S. Energy Information Administration.
- Vaillancourt, K. (2014, April). Electricity Transmission and Distribution. Energy Technology Systems Analysis Program E12, International Energy Agency.
- Vengosh, A., N. Warner, R. Jackson, and T. Darrah (2013). The effects of shale gas exploration and hydraulic fracturing on the quality of water resources in the united states. *Procedia Earth and Planetary Science* 7, 863–866.
- Von Meier, A. (2006). *Electric power systems: a conceptual introduction*. John Wiley & Sons.
- Vuong, Q. H. (1989). Ratio tests for model selection and non-nested hypotheses. *Econometrica* 57, 307–333.
- Wang, J., M. Shahidehpour, Z. Li, and A. Botterud (2009). Strategic generation capacity expansion planning with incomplete information. *IEEE Transactions on Power Systems* 24(2), 1002–1010.
- Wen, X., Y. Wei, and D. Huang (2012). Measuring contagion between energy market and stock market during financial crisis: a copula approach. *Energy Economics* 34, 1435–1446.
- Xunpeng, S. (2016). Gas and lng pricing and trading hub in east asia: An introduction. *Natural Gas Industry B* 3, 352–356.
- Yafimava, K. (2014). Outlook for the long term contracts in a globalizing market (focus on europe). Technical report, The Oxford Insitute for Energy Studies.
- Yanikoğlu, İ., B. L. Gorissen, and D. den Hertog (2019). A survey of adjustable robust optimization. *European Journal of Operational Research* 277(3), 799–813.
- Young, M. R. (1998). A minimax portfolio selection rule with linear programming solution. *Management science* 44, 673–683.
- Zeng, B. and L. Zhao (2013). Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operation Research Letters*, 41(5), 457–461.
- Zhang, H., J. Baeyens, J. Degève, and G. Cacères (2013). Concentrated solar power plants: Review and design methodology. *Renewable and Sustainable Energy Reviews* 22, 466–481.

BIBLIOGRAPHY

- Zhang, X. and A. J. Conejo (2017). Robust transmission expansion planning representing long-and short-term uncertainty. *IEEE Transactions on Power Systems* 33(2), 1329–1338.
- Zhang, X. and J. A. Conejo (2018). Robust transmission expansion planning representing long- and short-term uncertainty. *IEEE Transactions on Power Systems*, 33(2), 1329–1338.
- Zhao, L. and B. Zeng (2012). Robust unit commitment problem with demand response and wind energy. In *2012 IEEE power and energy society general meeting*, pp. 1–8. IEEE.
- Zugno, M. and J. A. Conejo (2015). A robust optimization approach to energy and reserve dispatch in electricity markets. *European Journal of Operational Research*, 247, 659–671.
- Zugno, M., J. M. Morales, and H. Madsen (2016). Commitment and dispatch of heat and power units via affinely adjustable robust optimization. *Computers & Operations Research* 75, 191–201.