Technische Universiteit Delft

Bidirectional softlinking for optimal energy planning

Luiscarlos A. Torres Sánchez

TU Delft

Bidirectional softlinking for optimal energy planning

by

Luiscarlos A. Torres Sánchez

to obtain the degree of Master of Science in Electrical Engineering

at the Delft University of Technology, the Netherlands.

To be publicly defended on Friday, September 18th, 2020 at 10:30 CEST.

Student number:	4805321	
Thesis committee:	Dr. ir. J. L. Rueda Torres,	TU Delft
	Dr. ir. M. Cvetković ,	TU Delft, supervisor
	Dr. ir. L. M. Ramírez Elizondo,	TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Abstract

The urgent transition to other forms of energy production, delivery and consumption has induced investments in new technologies such as renewable energy sources, storage units and demand response programs. Investments in these assets represent long-term commitments (e.g. 40 years ahead) from civil society, governmental institutions and capital investors. For the latter, traditional long-term investment models and their assumptions have been proven to be sufficiently accurate for decision making. Two core assumptions in these models are that electric demand profiles have recognizable patterns and that the generation side is able to fully follow them.

Nonetheless, with the introduction of the aforementioned technologies in the energy mix, these assumptions may become obsolete. For instance, renewable energy sources are intermittent, storage flattens peaks and valleys of demand curves and demand response introduces randomness in electric consumption. In comparison to long-term investments, however, these operations are much shorter in time (e.g. minutes, hours, days) and they are thus captured by the so-called power system models or short-term operations models.

In this context, this thesis evaluates possible effects that short-term operations may have on long-term investment decisions. To this end, this works first covers the models and computational tools typically used to define long-term investments and short-term operations in the electricity sector. Subsequently, a thorough research is conducted to explore possible methods that allow long-term investment models and short-term operations models to exchange relevant information to achieve a specific target. Among these methods, this work advocates for the bidirectional softlinking (BSL) due to the flexibility it offers to both models to be individually expanded and the ability to keep them apart, as independent entities. This thesis hence shows the potential of BSL, the challenges to adopt this mechanism and the guidelines for future research endeavors that could scale up the applicability of the method to actual power system planning with the inclusion of the aforesaid new technologies.

Preface

At commencement of this master's thesis, my enthusiasm for diverse topics encouraged me to try to reconcile them in one way or another. My supervisor, Miloŝ Cvetković, supported me in this idea, thus setting the start of an enjoyable, yet challenging journey. After weeks of asking all sorts of questions, I discovered an entire research area in power system planning that I subsequently grew fond of due to its relevance at this point in time. Now that I have finished, I can only look back and satisfactorily express that I have learned the most out of this project.

Within this context, I would like to thank Miloŝ for his technical support and the knowledge he shared with me during all these months. I would also like to thank José Rueda Torres, Laura Ramírez Elizondo and Peter Palensky for providing me with constructive feedback to deliver a high quality work. Furthermore, I would like to thank my employer, decon international GmbH, for financially supporting me throughout the two-year duration of the master's program. In particular, I thank Frederik Müller, Tobias Nentwig and Héctor Khalona Ramírez for their constant availability and help.

Lastly, I would like to thank my parents:

A ustedes les agradezco con todo mi corazón, pues siempre ha sido su amor y sacrificio lo que ha hecho posible que yo recorra nuevos senderos académicos y profesionales. A mi madre, quien cree en mí en todo momento; y a mi padre, quien siempre me motiva a cerrar carreras al estilo de Usain Bolt: "a toda potencia, a pesar del cansancio." Los amo.

Luiscarlos A. Torres Sánchez
The Hague, The Netherlands, September 2020

Contents

Nomenclature

Parameter	Variable	Index	Description	Explanation
R			Interest or discount rate	
N			Operational lifetime	
K			Annuity factor	$K = (1-(1+R)^{-N})/R$.
C^{cc}			Capital cost	Total overnight cost of building a power plant. Measurement: unit of cost per unit of power (e.g. kEUR/MW).
C^{acc}			Annualized capital cost	$C^{acc} = C^{cc}/K$. Measurement: unit of cost per unit of power (e.g. kEUR/MW).
C^{fix}			Fixed cost	Cost of maintenance, property taxes, insurances, etc. Measurement: unit of cost per unit of power (e.g. kEUR/MW).
C^{var}			Variable cost	Cost (e.g. fuel and carbon penalties) per electricity produced. Measurement: unit of cost per unit of energy (e.g. EUR/MWh).
C^{sup}			Start-up cost	Cost (e.g. extra fuel) for starting up a power plant. Measurement: unit of cost per unit of power (e.g. kEUR/MW).

P^{max}		Maximum active power	Measurement: unit of power (e.g. MW).
P^{min}		Minimum active power	Measurement: unit of power (e.g. MW).
P^{up}		Ramp-up limit	Measurement: unit of power (e.g. MW).
P^{down}		Ramp-down limit	Measurement: unit of power (e.g. MW).
d		Duration of an event	Measurement: unit of time (e.g. hour).
d_{minUp}		Minimum up time	Minimum time a generator must remain in operation. Measurement: unit of time (e.g. hours).
$d_{minDown}$		Minimum down time	Minimum time a generator must remain off. Measurement: unit of time (e.g. hours).
P^{d}		Power demand	Measurement: unit of power (e.g. MW).
I^{cap}	i^{cap}	Installed capacity	Measurement: unit of power (e.g. MW).
	t	Specific point in time	Measurement: unit of time (e.g. hour).
	c_{sys}^{total}	Total system cost	For example, sum of the total capital, fixed, variable, start-up and curtailment costs. Measurement: unit of cost (e.g. MEUR).
	c^{fac}	Capacity factor	
	p	Active power	Measurement: unit of power (e.g. MW).

u		Binary variable	Measurement: either 0 or 1.
e		Binary variable	Measurement: either 0 or 1.
s^{up}		Binary variable	Measurement: either 0 or 1.
g		Generator	
	b	Time block	It represents a time block with a specific duration (e.g. 1h, 3h, 5h, etc.).
	T	Optimization period	It represents an optimization period with a specific duration (e.g. half a year, one year, etc.). Moreover, each T contains a set of b's.
	H	Optimization horizon	It represents an optimization horizon with a specific duration (e.g. one, two, or multiple years). Moreover, each H contains a set of T's.

1
Motivation

1.1. Introduction

In 2018, countries around the world ejected a total of 55.6 billion CO_{2_eq} tonnes (i.e. 55.6 Gt) of greenhouse gas emissions to the atmosphere, with approximately 47% of them produced only by China (26%), the United States (13%) and the European Union (more than 8%) [1]. Estimations indicate that CO_2 emissions in particular reached in that same year 33.1 Gt-40.03 Gt, which represented between 1.7%-2% of increase in comparison to 2017 [1, 2]. While in fact emissions from all fossil fuels rose, the power sector accounted for nearly two-thirds of emissions growth (e.g. the use of coal in power alone surpassed 10 Gt of CO_2) [2]. With surging emissions and no signs of them peaking yet [1], the immediate restructuring of the electricity sector is deemed urgent to stop such environmental deterioration and its devastating effects on the climate.

This transition in the power sector has stimulated a wave of investment needs that seek to promote a new generation of technologies and an overall improvement of the system performance. The future electricity system is hence expected to include- among others- centralized generation such as large-scale wind farms, distributed generation such as solar-powered buildings, grid-scale and distributed storage, electric vehicles and demand side management [3]. In order to truly implement these changes, however, a *long-term* (e.g. up to 40 years in the future) commitment is usually required from civil society, governmental institutions and capital investors. For the latter in particular, these long-term investments represent crucial decisions given the considerable amount of time needed to obtain financial returns in such prolonged time frames. It is therefore at this point in time that experts in the electricity sector must test whether existing long-term investment models and their corresponding assumptions remain equally valid and insightful with the newly expected technologies.

Long-term investment models have traditionally considered that the generation side, mostly composed of conventional power plants, is almost always capable of responding to changes in demand. On the other hand, electric demand has been assumed to have steady patterns over the course of a given year; except perhaps, during some extreme weather conditions. For this reason, these investment models have regarded, for example, that six specific days with unique characteristics (e.g. coldest day in winter, hottest day in summer, a weekend day, a working day, etc.) are enough to fully represent electric consumption [4]. Based on these considerations, these models have provided long-term investment plans, which contain the best possible energy mix and the specific operation of each technology to supply demand in a certain geographical location.

Nonetheless, these two assumptions of a readily available generation and a predictable demand may no longer be sufficiently accurate for invesment planning. For instance, wind and solar power production are both intermittent and even show drastic intra-hour output changes [5]. This not only makes generation less reliable, but it also negatively affects the operation of conventional power plants that are exposed to a growing mechanical stress [6]. On the other hand, demand response programs that are expected to incentivize consumers according to the spot electricity price [7], will introduce a higher degree of randomness in electric consumption profiles. Another example in this regard is storage devices, which are expected to also act according to the spot electricity price, thus flattening both the peaks and valleys of power demand curves [3]. The actions implemented by these technologies are within a short period of time (e.g. minutes, hours or days) in comparison to the several years spanning the long-term investment decisions. For this reason, these and other actions from plausible technologies that fit within similar time frames are thus referred to as *short-term operations* or *short-term power system operations* in this thesis.

The relation between long-term investment decisions in these new technologies and their specific, yet unprecedented short-term operations, helped create the main question of this work:

How can short-term operations influence long-term investments in future electric power systems?

The target of this question is to consider the possible *future* in which technologies that are currently widely built and operated (e.g. conventional power plants) co-exist with a higher deployment of technologies on the rise (e.g. renewable energy sources and storage units). Such future is indeed a very concrete goal in Europe and the Netherlands (see the Dutch Climate Agreement in Ref. [8]). Moreover, the aim of this question is to obtain factual, numerical answers (e.g. total system costs, power system flexibility requirements, etc.) to measure the *influence* of the short-term operations in long-term investments. As a matter of fact, this entire work is crafted around these two central objectives. This chapter in particular provides deeper insights into *long-term investments* in Section 1.2 and *short-term operations* in Section 1.3. Thereafter, Section 1.4 introduces the stakeholders, Section 1.5 evaluates possible research gaps around this topic and Section 1.6 provides more precise, specific questions that delimit the research scope. This latter section also presents the methodology to be used in this thesis to approach the specific questions. Finally, this chapter closes with the outline of the rest of the document.

1.2. Long-term investments

1.2.1. Timescales, capital, risks and regulations

The primary function of a power system is to satisfy the energy demand of its end users. To this end, the system relies on an array of expensive technologies to produce and store energy and in a complex network infrastructure that interconnects those technologies with the points where energy is demanded. In the context of system design and expansion requirements, long-term investment frameworks can thus be consolidated in two stages: 1) the planning of what technologies will generate power to supply the demand; and 2) the planning of where these technologies will be located and the infrastructure needed to connect them to the existing system. The first step is a long-term projection referred as *generation expansion plan-*

ning, whereas the second step is a medium-term analysis called *geo-spatial planning* [9].

The generation expansion planning, usually published as a national *electricity master plan*, consists of investment decisions that span between 20 to 40 years or even more. These decisions represent political commitments and are often linked with long-term targets [9]. On the other hand, the geo-spatial planning addresses the site location of the generation projects and the financial aspects related to the corresponding transmission-line expansion required to incorporate the generating units in the grid. Some countries develop transmission plans that cover 5 years in advance, although others may extend the analysis to 15 or more years [9]. Fig. 1.1 shows a representation of these two studies in terms of their timescales. As a comparison, they are placed in relation to typical timescales of power system operations, which are further explained in Section 1.3.

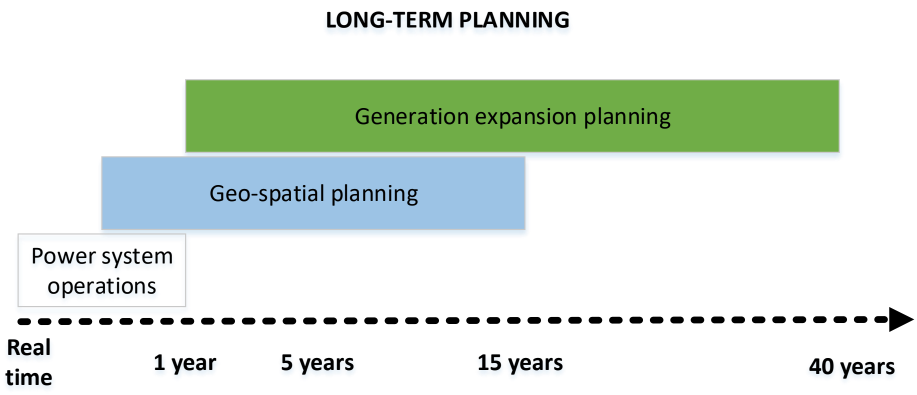


Figure 1.1: Stages in an electric power system design: 1) generation expansion planning (i.e. long-term investment decisions and political commitments) and 2) geo-spatial planning (i.e. location of the technologies to be built and their needs to be incorporated to the existing system). The timescales of these two stages are compared to the timescales of power system operations.

The technologies to be selected are not only costly, but also represent long-term investment decisions due to their typical technical lifetimes (e.g. a nuclear power plant may remain operational for 60 years, while a coal power plant for 40 years [10]). This implies that investors, whose main purpose is to obtain returns on their investments, must wait for a long time before collecting any financial benefits. This seemingly suggests that attracting investments in this sort of assets, that can neither be moved nor pay for itself for many years [11], represents a significant challenge. However, this in particular did not entail an impediment before the energy transition, as monopolies in the power sector were the main investors and their costs were simply charged to the consumers. In this way, investment decisions carried minimum financial risks and generation projects could therefore be financed through low interest rates [11].

Nevertheless, the energy transition brought along electricity market deregulation and increasing levels of renewable energy integration, which induced changing policy measures and market complexities [12]. This not only raised the interest rates for investments [11], but also made investors remarkably more aware of the potential risks to be faced under volatile conditions. For instance, these risks may include project-specific ones like an unexpected shorter technical lifetime of components [13]; or technical & management risks like the lack of sufficient local expertise to adequately operate or maintain new technologies [14].

Perhaps an even more relevant set of risks involves policy design risk, market design & regulatory risk and risk due to sudden policy changes [14]. To counteract these risks that are external to investors, governments play a fundamental role as it is their duty to incentivize investments by creating stable policy frameworks where retroactive policy changes are prohibitive; to communicate the value of these investments to society; to ensure clear and effective market signals; and to remove unnecessary regulatory barriers [12]. In fact, the role of the government has been so important in the power sector that over 95% of investments in 2018 were made by companies operating under fully regulated revenues or long-term contractual mechanisms to manage the risk of revenue associated with variable wholesale electricity prices [15]. This is because short-term electricity prices alone remain too low to trigger investments in the most capital-intensive assets, even in many countries with competitive wholesale markets [15].

1.2.2. Projected investments in emerging technologies

As covered in Subsection 1.2.1, the generation expansion planning represents a long-term investment strategy in a given set of energy technologies. In the European context, it is stipulated that the investments in these technologies will mostly be in onshore wind energy, followed by solar energy and gas power plants in the next decades [16, 17]. Furthermore, it is estimated that 950 GW of these sources need to be installed per year (equivalent to an annual investment of 80 billion EUR) to achieve high levels of renewable energy integration by 2050. Likewise, the investment costs needed in transmission and distribution reinforcement (i.e. geo-spatial planning in Subsection 1.2.1) to accommodate this newly-designed energy mix is approximately 105 and 443.7 billion EUR, respectively.

Furthermore, it is expected that between 43 GW to 90 GW of installed storage capacity will be required in the European Union by 2050 (with an investment estimate ranging between 80 billion USD and 130 billion USD, including the infrastructure needed for integration) [17, 18]. Similarly, investments on demand response are expected to increase due to the significant savings in network reinforcement and back-up power capacity that are obtained as a consequence. It is thus assumed that the investment requirement for the deployment of demand response programs and technologies will be about 770 billion EUR per decade until 2050 [16, 17].

In the Dutch context, the Ministry of Economic Affairs and Climate Policy issued the *Climate Agreement* [8] in 2019, which was drafted together with different energy sectors of the economy and who committed to performing fundamental changes in their ways of energy production and consumption. Although the expected levels of investments are not specified, it was agreed that renewable integration in the power sector should at least reach 70% (mostly wind and solar) by 2030. Furthermore, the vision for 2050 is to be absolutely carbon free in a scenario where sector coupling, demand response, cross-country interconnectors and storage play a fundamental role to alleviate the effects of a 100% renewable integration.

1.3. Short-term power system operations

As discussed in Subsection 1.2.2, the investment on gas power plants, renewable sources, demand response and energy storage is expected to increase in the next decades in the Netherlands and the whole European Union, under scenarios of realized high renewable integration. It is hence worth considering whether the short-term power system operations or more specifically, the technical constraints that these technologies have, their limitations, the interactions among themselves and the physical laws that govern electricity grids may directly influence this sort of long-term investment decisions. To tackle this question in greater detail, it is first necessary to delve into the definitions and categories found in these *short-term power system operations*.

Fig. 1.2 shows a representation of these operations, which have been grouped in three different categories based on how they are found in the literature: the grid (in reference to their original arrangement: network, transformers, conventional generators, etc.) [19], renewable energy sources [20] and storage units [21]. In the grid, it is possible to find ultra-fast phenomena such as lightning propagation and switching surges. Generally, the former studies the effects of lightning strikes on the power infrastructure, while the latter models the effects of fast deployment of electrical equipment.

Closer to the seconds scale, the grid may face transient instability due to line faults, severe frequency drops due to the loss of a generator or due to the lack of enough inertia in the system to damp disturbances. At this time scale, fast storage units such as batteries may be able to provide solutions for these problems by providing the needed operational inertia and a fast frequency response. From the seconds up to the minutes, the power system status can be checked via a power flow calculation. From the results, it is possible to observe whether the grid is experiencing voltage instability, for instance, after a line fault has been cleared. Additionally, renewable energy sources may change their output drastically due to steep variations in the weather conditions, while storage units may alleviate the effects of this uncertainty by providing power reserves.

From the minutes to the hours and all the way to a few days, the influence of the renewable sources becomes more predominant in the power system operations. First, forecast errors arise as the mismatch between actual and projected outputs becomes evident. In order to correct for this behavior, storage units may be deployed (i.e. follow the load), while online conventional power generators may readjust their output; depending on what series of actions represents the least cost at the moment. Finding that least, optimal cost, is implemented via an optimization problem called *economic dispatch*. Second, an aspect with a longer-term effect on the system is the variability in the load produced by the renewable energy sources. As the power generated by these sources must usually be consumed (in a clear analogy to the energy demand that must always be supplied), they tend to be interpreted as a negative load. Therefore, their addition to the load profile itself is known as the *net load* and it shows a more variable behavior that must be absorbed by the supplying technologies.

From the set of suppliers, power plants need to be carefully scheduled while considering the aforementioned net load variability. Their schedule needs to be decided one or a few days in advance before the day of energy delivery, because of their many technical constraints or due to financial reasons. Among the technical constraints, some technologies must remain off for several hours if they are to be shut down (e.g. a nuclear power plant may be required to remain off for up to 168 hours [10]); while, on the opposite end, they must also remain on for a certain amount of time if they are to enter in operation. On the financial side, these power plants have significant costs when starting up as they normally need much more fuel to move their heavy machinery. Thus, with several generating units in the system and multiple constraints and costs to consider, the problem becomes another optimization to find the least cost for delivering energy demand, known as *unit commitment*. The solution to this problem then gives the exact schedule of the power plants to be online within the next day(s) and at exactly what hour. Finally, the remaining power system operation refers to seasonal storage, which exemplifies that there exist other sources of flexibility on the supply side such as the hydro-reservoirs that are able to store, for instance, the overproduction of solar energy in the summer to subsequently release it during high load demand peaks in winter.

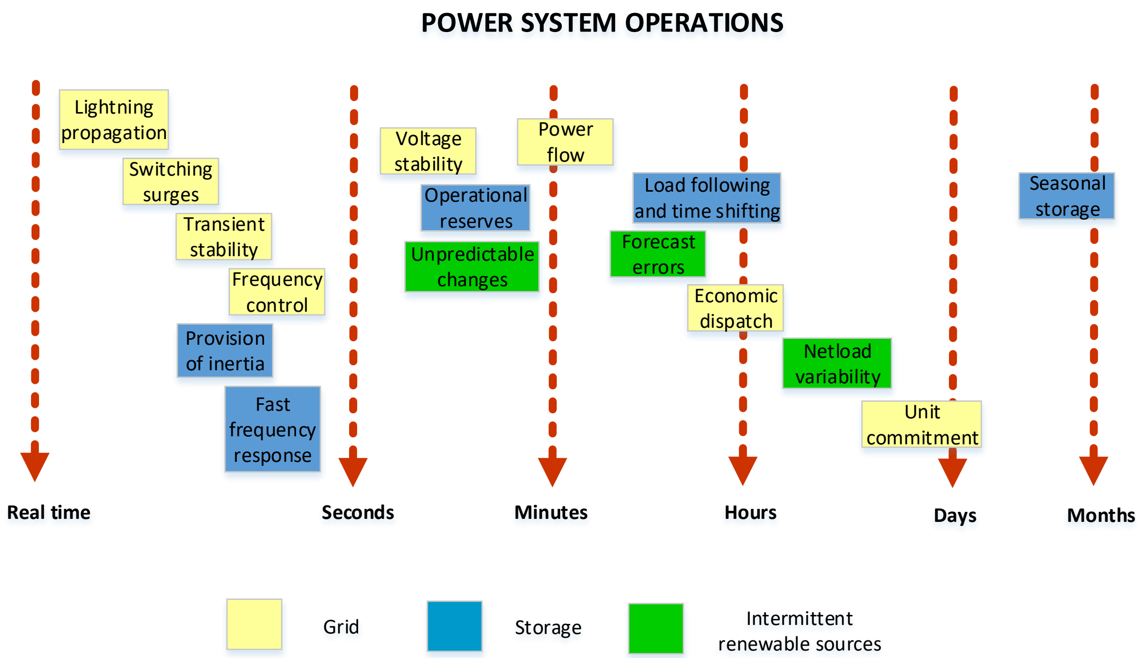


Figure 1.2: Power system operations representing the grid (i.e. a traditional power system with conventional generators, transformers, network infrastructure, etc.), intermittent renewable energy sources and energy storage units. The operations and their corresponding timescales to create this graph have been compiled from three different sources: grid [19], renewable energy sources [20] and energy storage units [21].

1.4. Stakeholders

The central question in this thesis so far has been crafted around the idea that short-term power system operations, as described in Section 1.3, can potentially influence the sort of long-term investment decisions, introduced in Section 1.2. To further explore this research question, it is crucial to identify *who* would be interested and *why* the potential answers would be beneficial to them.

1.4.1. Who are they?

- **Investors:** In the power sector, it is possible to find retail, institutional and utilities investors [22]; although banks (e.g. European Investment Bank) are also mentioned as one of the key project financing providers in OECD countries [23]. Retail investors are individuals that place their resources in renewable energies based on their own initiative; while institutional investors, such as pension funds, insurance companies, investment funds, etc. manage their own (or third-party) investments [22]. On the other hand, utility companies are investors that are responsible for the generation, transmission and distribution of electricity as a part of their core business. In Germany, for example, retail investors engage in renewable energy projects with at most 1 MW in size, whereas institutional investors with 13.4 MW on average for wind projects. In contrast, utility companies show much more interest in conventional generation than in renewable capacity [22].

- **Energy policy makers:** For a specific country, energy policy can be analytically distinguished between its *external* and its *internal* side. The external side corresponds to the decisions taken to guarantee security of energy supply in the country while considering the decisions of other states, the geo-political circumstances, etc. [24]. The main actors of these decisions are the national governments (i.e. the head of the state in particular, accompanied by the ministers of foreign affairs), but also large energy firms that are directly influenced by national executives; or even diplomats and experts in international relations as in the case of the European Union [24]. On the other hand, the internal side is associated with the decisions related to the use of energy within the national territory (production, transport, distribution, sale, energy saving, etc.). The main actor on this side is again the central government (e.g. through the ministry of energy in some countries, or simply other ministries engaged in these activities) [24]. Other actors found on the internal side are firms that are active in the energy sector and that have formed various organizations to represent their interests and exert an influence on the energy policy making process. Moreover, due to more market-oriented sectors, it is possible to find independent authorities, who undertake a set of tasks that were previously performed by the government, such as the determination of prices, as well as technical and quality standards, and market control. Other various actors on the internal side are local governments, large industrial clients and groups of experts [24].

- **Grid operators:** This typically refers to transmission system operators (TSOs). Some of the core activities of these operators include maintaining an adequate network development plan, renewal and maintenance of existing network components, defining common technical requirements for a secure system operation, implementing common procedures for congestion management in the grid, managing disturbances and activating remedial actions, ensuring system balance and enough transmission capacity, etc [25].

1.4.2. Why would they be interested?

Investors on conventional power plants would be interested in incorporating the effects of short-term power system operations on their investment models, as this time resolution allows them to include cycling requirements (e.g. generator minimum time and start-up costs, covered in Section 1.3). With this information in hand, it is therefore possible to have a better understanding of the plant's financial depreciation and technical wear and tear over time. Furthermore, understanding these requirements together with the behavior of other components in the grid would allow them to better calculate the expected energy to be sold in the wholesale market. These considerations become particularly important with increasing levels of renewable energy integration, since these sources not only force power plants to cycle more often, but also significantly reduce the plants' generated energy due to the current existing market rules in which renewable energy in-feed is prioritized.

On the other hand, renewable energy investors would also be benefited by having a more specific short-term information on their investment decisions. An example of this is the reduction of their energy curtailment by the system operator. This can be implemented by understanding the system integration capabilities at this timescale (e.g. flexibility, network flows, power demand profiles, etc.). On the contrary, not intending to reduce the impact of this risk can remarkably affect expected project revenues. For instance, Ref. [26] indicates that, although there exist Power Purchase Agreements (PPAs) that fully compensate for the energy curtailed and its respective loss of green credits, these agreements are in fact rare so that curtailment almost always entails a negative investment risk. For example, in the case of wind energy in the U.S.A., some operators do not compensate for curtailment at all, whereas others do not compensate if curtailment is due to reliability issues. In other countries, renewable energy generators are compensated for only a fraction of the energy lost [26].

In the case of energy policy makers, they need to understand how their policies and measures *truly* incentivize investors. This directly translates to understanding how investors take their long-term investment decisions, while including the influence of short-term operations (e.g. the case of power plant investors). Furthermore, the design and implementation of certain energy policies can be greatly enhanced by using similar techniques. For instance, demand-side management requires an understanding of customer behavior, utility and elasticity, as well as fluctuations of spot prices [7]. All these aspects can vary significantly depending on the hour and the day [27], and not only on longer time periods, such as seasons. Furthermore, if energy policy makers do not include short-term flexibility requirements to draft their carbon policies, it may occur that strict carbon regulations prompt a shift toward low-carbon baseload generation (e.g. nuclear) that is inflexible [28], leaving grid operators with a vulnerable fleet to react to fast changes in grid operations.

For energy storage investors, short-term information is of paramount importance to their business models. As detailed in Section 1.3, storage units are capable of providing five system services to the electricity grid; among which, four of them (i.e. provision of inertia, fast frequency response, operational reserves, load following and time shifting) lie within the short-term timescales. Thus, failing to consider this time frame would surely provide inaccurate information on the potential project revenues. Finally, the motivation of grid operators can be reflected in terms of costs and system security. For instance, if conventional power plants are the main source of flexibility in a given system, the lack of power plant intertemporal constraints may not only indicate false levels of security [29], but also misrepresent the costs for balancing the grid. Both aspects become of increasing importance with the rise of renewable energy integration.

1.5. Research gaps and contributions of this thesis

The very fact that the main question in Section 1.1 intends to relate long-term investments to short-term operations indicates that these two fields have been traditionally examined separately. More specifically, energy planners have been in charge of analyzing long-term investment decisions, while power system engineers have conducted the corresponding studies to ensure that these decisions are technically realizable in a given system. Nonetheless, most of the existing literature that explores these two fields in an integrated manner seems to lack the necessary background that would allow either expert to understand how these two engineering areas are entangled and how they can be modeled in general. This thus constitutes the first research gap addressed in this thesis, which is filled in by providing the reader with specific guidelines from an upper level perspective (i.e. investments) to a lower level one (i.e. operations), including how they are connected and why they should be considered.

Furthermore, most of the literature found also offers little to no detail about how these two models can exchange information in general (i.e. other than the method proposed by each individual article). This thesis thus seeks to tackle this second research gap by incorporating different methods in order to compare them and analyze them, in such a way that any given expert is able to take an informed decision or know where to find additional information. Finally, one of the recent methods proposed in the literature, the *bidirectional softlinking* (BSL), which is also the focus of this thesis, seems to lack rigorosity in terms of its implementation, classification and explanation in most studies. This thesis therefore intends to fill in this third research gap by providing mathematical and analytical arguments about the BSL method, while laying the foundations for more in-depth future research works.

1.6. Specific research questions and the methodology to approach them

The main question posed in Section 1.1 leads to three specific research questions that are thoroughly assessed in the next chapters based on the following methodology:

1. How can long-term investments and short-term operations in the power system be modeled and exchange *relevant* information?

   The first part to answer this question is to understand what *long-term* investments are with respect to energy systems (as a bigger part of the electricity sector). This entails finding examples of how researchers approach their investment planning, the scenarios they use, the questions they intend to solve and whether their approach is applicable to this thesis. Subsequently, the research is focused on the long-term investment models of the power system itself. To this end, several comprehensive reviews with their corresponding comparisons between these models are evaluated. After this task, a specific model is selected based on its advantages over other models to represent the electricity sector. Later, a deeper assessment is required to unveil the model's limitations.

   The second part to answer this question is to identify what the best way to model short-term operations is. In particular, find what kinds of models there are to do so and which of them could have a stronger influence on the previously chosen long-term investment model. Afterwards, it is necessary to select specific, validated computational tools that recreate both the long-term investment and the short-term operations models. The third part to answer this question consists in researching several literature works that propose diverse ways in which these two models can exchange information. Once these ways have been categorized, a specific one has to be selected for further assessment. Finally, a very crucial step in this question is to define what the *relevant* information to be exchanged is in the first place. This has to be determined after several tests in which the most representative parameters and variables from each model become evident. Once this is done, it is possible to understand how they relate to, and therefore how they influence each other.

2. Under what circumstances do short-term operations become influential on long-term investments in the power system?

   A set of case studies is developed to approach this question. These case studies intend to recreate the combined circumstances of a traditional power system with those of a future (e.g. up to 40 years ahead) power system. In the first case study, the long-term investment model decides how to distribute the installed capacities of two power plants. Subsequently, these power plants are assigned more detailed equations, which help represent their corresponding short-term operations. For this case study, the circumstances consists of different levels of flexibility for the short-term operations equations. Moreover, a considerable amount of renewable energy integration is assumed. In general, the first case study also works as a benchmark to explain and elaborate about the mechanisms used to exchange information between the long-term investment model and the short-term operations model. The specific mechanism to be thoroughly explained and justified is the *bidirectional softlinking*, as aforementioned.

   Thereafter, the second case study intends to test the scalability of the methods developed in the first case study by adding two more power plants. This case study then also adds more circumstances since the inclusion of these new power plants brings a different family of short-term operations. Furthermore, a new, higher level of renewable energy integration is incorporated. Finally, the third case study considers the entire Dutch fleet of power plants, while the predominant circumstances are the curtailment of renewable sources and batteries performing arbitrage for economic benefits. In this way, the last case study brings in circumstances expected in future power systems.

3. How impactful are the changes produced by short-term operations in comparison to the scale of long-term investments in the power system?

   After all case studies have been implemented, a numerical comparison will be performed to assess the impact from short-term operations and the circumstances considered in each case study. For instance, if a given variable is selected to be monitored (e.g. total system cost), its value prior and after the model coupling will be compared. This would act as a metric to assess whether the bidirectional softlinking is able to bring potential benefits to the overall costs, system performance, etc.

1.7. Outline of this thesis

To answer the main question posed in Section 1.1, the rest of Chapter 1 has covered different aspects regarding *long-term investments* and *short-term operations* in the power system. This includes definitions, timescales of their corresponding processes and the stakeholders that might be interested. Building on these ideas, this chapter also indicates the specific research questions of this thesis. Hereafter, Chapter 2 elaborates on the mathematical and computational models used to study long-term investments and short-term power system operations independently. That chapter also presents the advantages and limitations of these models, as well as an exhaustive assessment of their attributes based on the author's experience when dealing with them. Chapter 3 explains three mechanisms in which these independent models can exchange information to improve the overall system costs and performance. Moreover, each of these mechanisms is accompanied by a detailed literature review of scientific works that corroborate their respective methodologies. In this thesis, the focus is placed on a recently suggested approach called *bidirectional softlinking*. Chapter 4 then exemplifies the use of an algorithm to bypass obstacles in long-term investment models, while it also poses three specific case studies to test the scalability and areas of application of *bidirectional softlinking*. Finally, Chapter 5 discusses the results obtained, formulates conclusions and provides recommendations for future works on this topic.

2

Models: Implementation and selection

2.1. Introduction

In Chapter 1, long-term investment decisions in the electricity sector as well as their corresponding characteristics were introduced. This chapter thus further elaborates on the different ways in which these decisions can be represented by using the so-called *energy system models*, while taking into account the influence that several energy sectors of the economy exert on each other. In addition, Chapter 1 introduced the nature and types of short-term power system operations and how their time duration relates to the timescale of long-term investment decisions. In this chapter, particular attention is paid to the model and mathematical formulation of the unit commitment and economic dispatch (UCED) operations, as they are argued to have the most significant influence over investment outcomes. Finally, this chapter discusses the limitations of each model and how they can potentially be solved by allowing them to exchange information.

2.2. Energy system models

2.2.1. Examples of application

Energy system models are used to assess the resulting effects of changing policies, pollution constraints and technological development trends on long-term investment decisions. For instance, Ref. [30] uses a high geographical resolution of China to discuss three potential decarbonization pathways in four *coupled* energy sectors (electricity, transportation, heat and industry) based on relaxed to more stringent constraints on CO_2 emission budgets. A particularly important contribution of this article is to remark that the target of the Chinese government to reach the highest peak of emissions in 2030 is not ambitious enough to maintain the average global temperature rise below 2C, given that China is the largest greenhouse gas emitter in the world [30]. Similarly, Ref. [31] discusses the possible decarbonization pathways for the U.S. for 2050. Based on four scenarios that are built from different combinations among five crucial elements (e.g. carbon-capture systems availability application, fuel switching, etc.), this report finds that it is technically feasible to achieve 80% greenhouse gas reduction by 2050 in the country. The *expected* incremental cost to achieve this goal is 1% of the projected 40 trillion USD U.S. gross domestic product (GDP) in 2050 [31].

Energy models can also be easily increased in scale to include larger regions that are geographically interlinked. For example, a case study about Latin America can be found in Ref. [32] for the years 2010-2050, where it is specified that the region is expected to double their annual investment in the energy sector with additional 330 billion USD in case that targeted climate policies are implemented. Although this increment is considerable in absolute terms, such investments actually decline over time in relation to the GDP [32]. A second application example can be found in Ref. [33] where Africa's 2050 energy outlook is discussed. This article concludes that Africa is capable of launching energy systems that rely on renewable sources from the onset rather than following the high-contamination patterns of the West [33], but only if stringent climate policies are implemented. Regardless of this positive outcome, Ref. [33] claims that the goals set by the Africa Renewable Energy Initiative (AREI) to deploy 300 GW of renewable power by 2030 is probably unrealistic.

These models may also play a role in understanding the effect of the energy transition in different aspects of society or how certain measures may help accelerate such transition. For instance, Ref. [34] discusses the problem of stranded assets, which usually refers to conventional generators that must face a premature devaluation [34]. This article concludes that a substantial amount of stranded capacity is to be expected in the medium term in Europe if the region is to seriously adhere to the climate targets recently established. This holds even when investors have perfect foresight over the development of the markets, but since energy systems are linked to energy policies, which are intrinsically imperfect in foresight due to political decisions, the number of stranded assets in Europe will be further aggravated [34]. Another interesting study in Ref. [35] analyzes the impact of the electricity supply decarbonization on land requirements. Using fixed carbon emission caps, Ref. [35] indicates that the increasing share of wind and solar power will have a ten-fold impact increment by 2030 in comparison to the current energy system requirements in Alberta, Canada. Therefore, it is essential that land requirements are also included as differentiating factors when choosing between decarbonization alternatives [35]. Lastly, Ref. [36] studies the long-term effect of demand response on the Portuguese energy system. After having established three different scenarios, authors state that demand response is able to reduce overall costs, which include a decrease in the capital cost due to a diminished investment on new capacity and the cost of system operation (therefore, lower electricity price) since more renewable power will be available at a zero marginal cost [36].

2.2.2. Classification

The classification of energy systems may be traced back as early as the seventies when the first surveys of global and international energy models appeared [37]. Currently, several review and categorization works [38–43] tend to indicate three key contributions in this field: van Beeck [44], Jebaraj-Iniyan [45] and Bhattacharyya-Timilsina [46]. In this context, Ref. [41] merges these three classifications with other works in order to accommodate newer approaches that have been proposed for energy systems modeling in recent years such as neural networks, agent-based and fuzzy logic.

Since the classification scheme proposed in Ref. [41] constitutes a detailed and extensive framework, it is adopted as the standard for this thesis and it is thus subsequently interpreted and explained in Table 2.1. For a richer description of each aspect, the reader is referred to these sources where the information was compiled and presented, i.e. Refs. [41, 44–46].

Table 2.1: Classification of energy system models based on previously defined categories and new categories. The classification is based on the work of Ref. [41], with direct influence from [44–46].

Category	Sub-category	Specification	Explanation
1. Purpose of the model	General	Forecasting	Prediction based on extrapolation of trends found in historical data.
		Exploring	Analysis of future scenarios (i.e. comparison of alternative ones with respect to business as usual).
		Backcasting	Creation of envisioned future scenarios and the current decisions that may be needed to achieve them.
	Specific	Energy demand model	The focus lies on the demand, which may be a function of changes in population, income, energy prices, etc.
		Energy supply model	The focus lies on the supply. That is, how to best meet demand in technical and financial terms?
		Impact model	Assessment of the consequences of certain policy measures or changes in the financial, social or environmental situation.
		Appraisal model	Comparison and appraisal of several options in order to select the most suitable one.
2. Structure of the model		Degree of endogenization	Extent to which parameters within the model equations are native to the model and are not externally incorporated.
		Description of non-energy sectors	Extent to which non-energy sectors (e.g. investment, trade, income distribution, etc.) are detailed in the model.
		Description of end-uses	Extent to which end-uses are modeled. For instance, the more details about end-uses, the better energy efficiency measures can be studied.
		Description of supply technologies	Extent to which technologies can be described. For instance, some models may treat technologies as a blackbox, which may not allow for a clear outcome of the optimal energy mix.
3. Geographical coverage		Global	It reflects world aspects (e.g. global economy).

		Regional	It refers to continental regions such as Latin America, South Asia and Europe.
		National	It refers to countries. Nations treat world market conditions as exogenous, but simultaneously consider inter-dependencies between major sectors within their territory.
		Local	It refers to sub-national territory; that is, regions within a country.
		Single project	It refers to a particular site, that is usually at sub-national level.
4. Sectoral coverage		Single-sector	It provides information on a particular sector (e.g. energy sector) and does not consider the macro-economic linkages of that sector with the rest of the economy.
		Multi-sector	It can be used at the international, national, as well as sub-national level to focus on the interactions between these sectors.
5. Time horizon		Short	5 years or fewer.
		Medium	Between 5 and 15 years.
		Long-term	15 years or more.
6. Time step		Minutes, hours, monthly, yearly, five-yearly, user-defined.	
7. Renewable technology inclusion		Hydro, solar (PV and thermal), geothermal, wind, wave, biomass, tidal, etc.	
8. Storage technology inclusion		Pumped-hydro energy storage, battery energy storage, compressed-air energy storage, hydrogen production, storage and consumption, etc.	

9. Demand characteristic inclusion	Transport	Internal-combustion vehicles, battery EVs, vehicle-to-grid EVs, hydrogen vehicles, hybrid vehicles, rail, aviation, etc.	
	Residential	Heating, lighting, cooking, appliance usage, smart appliances and smart meters, etc.	
	Commercial	Offices, warehousing, retail, etc.	
	Agricultural	-	
10. Cost inclusion		Fuel prices, fuel handling, investment, fixed operation and maintenance (O&M), variable operation and maintenance (O&M), CO2 cost, etc.	
11. Analytical approach		Top-down	It follows an *economic approach*. For instance, it regards technology as a set of techniques by which inputs such as capital, labor, and energy can be transferred into useful outputs. Therefore, an economic approach does not have a detailed representation of technologies.
		Bottom-up	It follows an *engineering approach*. It describes the techniques, the performances, and the direct costs of all technological options in order to identify possibilities for improvement.
		Hybrid	It introduces technological detail within a macro-economic approach. These can be of two forms: a coupling of existing bottom-up and top-down models ('soft-linked'), or a single integrated model, which combines both features.
		Other	

12. Underlying methodology		Econometric, macro-economic, micro-economic, economic equilibrium, optimization, simulation, stochastic/Monte-Carlo, Spatial(GIS), Spreadsheet/toolbox, accounting, etc.	
13. Mathematical approach		Linear programming, mixed-integer programming, dynamic programming, fuzzy logic, agent-based programming, etc.	
14. Data requirements		Quantitative	It can be used when very detailed outcomes are required for a certain data set.
		Qualitative	It can be used when data are not available or are unreliable.
		Monetary	
		Aggregated	Highly aggregated data with little technological detail.
		Disaggregated	Great details provided by diasaggregated, specific data.

2.2.3. ESOMs: Central generation expansion planning (CGEP)

For this thesis, it has been decided that the energy model to be selected must follow an engineering approach that allows to have a rich representation of the technologies involved, as well as a clear evolution of their costs over time. Thus, bottom-up models, and in particular, the group of *energy system optimization models* (ESOMs) has been chosen for this purpose. ESOMs identify the solution of the energy system by computing the investments and operation that result in an equilibrium between production and consumption of commodities [47]. The equilibrium is only *partial*, because ESOMs consider only energy systems (power, transport, heating sector, etc.), which are in fact a part of larger, more influential systems such a national or global economies [47] that can easily disturb such equilibrium.

When the power sector is of interest, engineers tend to employ a sub-model of ESOMs [48] known as the *central generation expansion planning* (CGEP) problem. The CGEP relies on the idea of a government-regulated central planner who is benevolent. As such, it intends to maximize social welfare [49, 50] and simultanously create incentives for private investors to direct their resources to the electricity sector in liberalized markets [51]. When demand is inelastic, maximizing the social welfare is equivalent to minimizing the costs of electricity production and delivery [51]. For this reason, the CGEP problem is usually formulated in such a way that multiple costs (e.g. investment, outage and emission costs) are merged into a single-objective function, whereas other objectives such as specific renewable targets or flexibility requirements are inserted as constraints [52]. Other approaches incorporate instead multiple objective functions such as the minimization of environmental pollution, nuclear hazards, portfolio investment risk or maximization of budget, renewable penetration level, etc. in order to provide decision makers with insights into potentially conflicting aims [52].

In relation to the market, the CGEP assumes that all agents possess perfect information and therefore play under perfect competition [49, 50]. The idea of perfect competition tends to be considered unrealistic [49, 50] and in fact provides completely different results when compared to those obtained from actual power system data (see Germany's case [53]). Nonetheless, it is also claimed [53, 54] that the increasing use of interconnectors with sufficient transmission capacity will eventually lead to near-perfect or truly-competitive electricity markets. The assumption of perfect competition has two important consequences: 1) the minimization of costs is equivalent to the profit maximization of each of the agents [49, 50] and 2) none of the agents is able to influence the price, then they all behave as price takers [53] so it can be mathematically proven (see Ref. [55]) that the resulting electricity price is equal to the marginal production costs (the marginal cost of the most expensive operating unit that has not yet reached its power output limit); a principle on which the so-called *merit-order* is based on.

CGEP: Single-period and multi-period optimization

The central generation expansion planning (CGEP) problem can be categorized based on how investment and operation decisions are taken during a given optimization horizon. For instance, consider the planning horizons represented in Fig. 2.1, where each of them is subdivided in planning periods (p.p.). The *static model* corresponds to a single-period optimization that takes the CGEP decision (i.e. D1) at the onset of the horizon [52], while typically considering the system characteristics that are expected at the end of the horizon [51]. This is a practical approach for systems that are not anticipated to undergo drastic changes over the horizon, but instead predictable ones like an increase in demand levels. Although this method is simple, it can lead to over-investment on capacity that may only be needed towards the end of the horizon [51]. Therefore, a more accurate approach is the *dynamic model*, which corresponds to a multi-period optimization that takes the best possible sequential CGEP decisions (i.e. D1, D2, ..., Dk) at each period so that the optimal investment on technologies can be implemented at different times during the planning horizon. As observed, this approach is increasingly more flexible as it allows to incorporate expected changes in the demand, oil, gas and coal prices (highly volatile throughout the years) over the time frame considered [51].

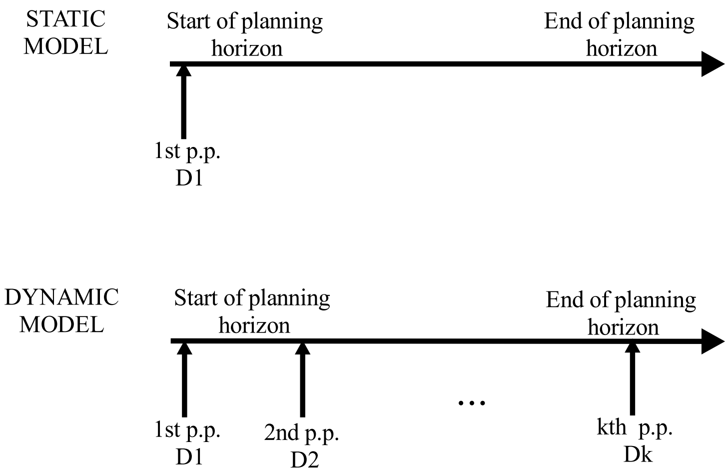


Figure 2.1: Static and dynamic central generation expansion planning (the figure has been adapted from Refs. [51, 52]). The static method consists of a single decision while taking into consideration the expected characteristics at the end of the horizon. On the other hand, the dynamic method allows for multiple sequential decisions along the horizon that can better represent changing external parameters.

Single-period optimization (static model):

The static model minimizes the total system cost c_{sys}^{total} in a given period T. For a conventional energy mix and an inelastic demand, the simplest form of the single-period optimization can be represented as follows:

$$\min_{i_g^{cap}, p_{g,b}} \quad c_{sys}^{total} = \sum_g \left(\left(C_g^{acc} + C_g^{fix} \right) i_g^{cap} + \sum_b C_g^{var} p_{g,b} \cdot d_b \right) \tag{2.1a}$$

$$\text{s.t.} \quad 0 \leq p_{g,b} \leq i_g^{cap} \tag{2.1b}$$

$$i_g^{cap} = e_g P_g^{max} \tag{2.1c}$$

$$\sum_g p_{g,b} = P_b^d \tag{2.1d}$$

$$b \subseteq T \tag{2.1e}$$

$$T \subseteq H. \tag{2.1f}$$

As observed, the only relevant parameter of the horizon H is the single period T. The lack of multi-period equations does not allow to assess the evolution of capital costs C^{cc} or asset depreciation over time; thus, in order for technologies to compete on the same basis, they must be compared through their annualized costs C^{acc}. With regards to the mathematical formulation, the binary variable e_g in Eq. (2.1c) gives the problem the ability to build power plants with discrete sizes, which is how they are usually offered by manufacturers. The existence of e_g turns the problem into a mixed-integer linear optimization, which increases the computational complexity of the solution and erases its convexity. If i_g^{cap} is considered as a continuous variable and therefore Eq. (2.1c) is dropped, the problem can be solved by *purely* linear optimization techniques or graphically, by a method called *screen curve*.

Standard screening curve

While the *screen curve* may provide elegant solutions that try to capture different aspects (e.g. thermal cycling [56], ancillary services [57], existing capacity [58] and planned outages [59]), the most commonly used version is the standard screen curve (SSC), introduced in Ref. [60]. The SSC is a green field methodology, which means that the focus is on a single year and no previous capacity is considered [61]. Therefore, the optimal generation mix is as "built from scratch" [49].

The standard screen curve (SSC) method compares the generating costs per unit of power c_g^{gcp} of conventional power plants, which are represented by time-dependent affine functions that increase with utilization time:

$$c_g^{gcp}(t) = C_g^{acc} + C_g^{fix} + C_g^{var} \cdot t. \tag{2.2}$$

When these functions are compared, they typically intersect due to the presence of baseload power plants (e.g. nuclear) that are expensive to build and maintain, but have low variable costs; load-following power plants (e.g. coal) with relatively balanced costs of construction, maintenance and variable operation; and peakers (e.g. combined-cycle gas turbine- CCGT) and high-peakers (e.g. combustion open-cycle gas turbine- COCGT), which are cheap to build, but are extremely expensive under variable deployment. These intercepts form a piece-wise function that represents the minimum possible total generating cost per unit of power of the energy mix.

After the intercepts have been found, they are mapped to a sorted demand profile (e.g. load duration curve, LDC [1]), which finally provides the optimal installed capacity per technology when the abscissa coordinates are projected onto their corresponding points on the vertical axis. To better depict the concept, Fig. 2.2 compares four power plants with characteristics as specified in Table 2.2. The intercepts occur at $t = 1350.359$ h, $t = 3792.934$ h and $t = 8283.327$ h, which once mapped to the LDC, provide each technology's installed capacity: $i_{nuc}^{cap} = 531.781$ MW, $i_{coal}^{cap} = 188.417$ MW, $i_{ccgt}^{cap} = 94.024$ MW, $i_{occgt}^{cap} = 185.778$ MW. With slightly more effort, the theoretical capacity factor per technology can also be obtained: $c_{nuc}^{fac} = 0.998$, $c_{coal}^{fac} = 0.664$, $c_{ccgt}^{fac} = 0.276$, $c_{cocgt}^{fac} = 0.051$. This ultimately allows to calculate the resulting total system cost: $c_{sys}^{total} = \sum_g c_g^{gcp}\left(t = c_g^{fac} \cdot 8760 \cdot 10^{-3}\right) \cdot i_g^{cap} = 255.976$ MEUR, based on the units in which variables have been considered in Table 2.2.

In some cases, a virtual technology with zero capital and fixed cost, and an extremely high variable cost (e.g. set to the value of lost load, VOLL, estimated at 22.94 kEUR/MWh in the Netherlands [63]) is also added to

[1] When renewable power is substracted from the load, the term used is either *net load* or *residual load*, and their sorted profiles are the net load duration curve (NLDC) and the residual load duration curve (RLDC), respectively. In the literature, these two terms are employed interchangebly in several cases, and although it seems that there is no confusion with regards to *net load*, *residual load* may also include the substraction of the must-run requirements (for instance, in Ref. [62]). To avoid any misunderstanding, this thesis adheres to the term *net load*.

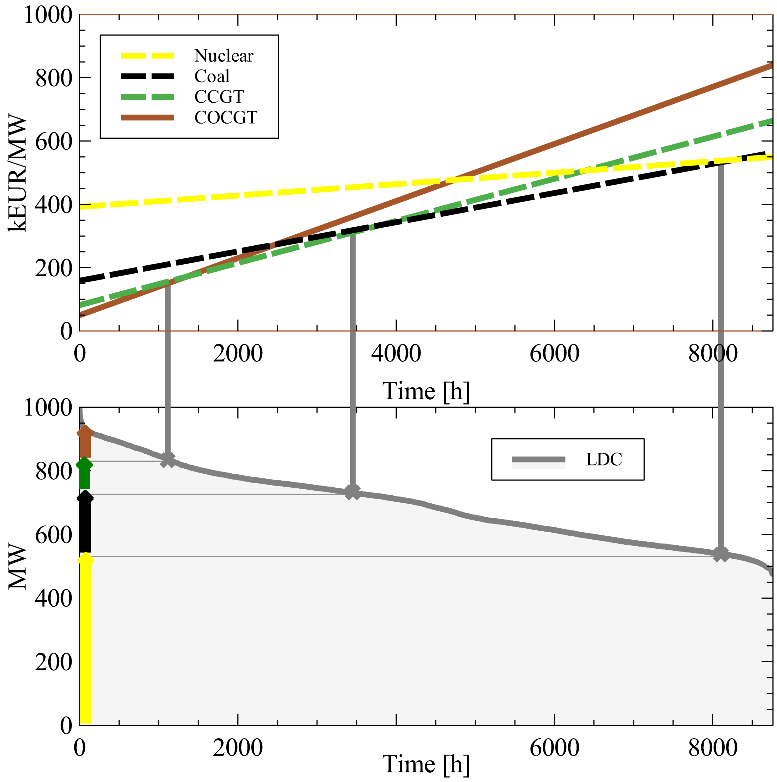


Figure 2.2: Standard screen curve (SSC) method in which intercepts of the generating costs per unit of power c_g^{gcp} are mapped to a load duration curve (LDC) to obtain the optimal generation mix. The LDC here shown is a re-scaled version of the 2018 Dutch demand profile, obtained from the ENTOSE transparency database.

the standard screen curve. The intercept that this virtual technology produces is also mapped to the LDC and represents those hours in which it is cheaper to curtail load than supplying it. The interested reader is referred to [49, 50] for a richer explanation about the role of this virtual technology in the current research debate between energy-only markets and capacity-based markets.

As observed, the SSC is a simple method and as such, it suffers from severe shortcomings in actual generation planning studies. Although some of these disadvantages have been addressed as aforementioned ([56–59]), multiple still remain. For example, SSC cannot consider renewable sources as an invesment choice [64], power plant decommissioning and unplanned outages are disregarded [57]; and generating units are misrepresented as they cannot be chosen on discrete sizes [64] and are either operated at full output or are shut down [65]. Nonetheless, this method is still widely implemented due to its simplicity and most importantly, because it provides the modeler with a first approximation of the system under study [66].

Table 2.2: Financial costs and technical parameters of four different technologies: nuclear, coal, combined-cycle gas turbine and combustion open-cycle gas turbine. Capital costs, variable costs (exchange rate 1.333 USD per EUR) and operational lifetimes are obtained from Ref. [67]. Fixed costs are obtained from Ref. [10]. The R=0.07 discount rate has been selected as it approximately corresponds to the market rate in deregulated or restructured markets [67]; which fits the European context.

Technology	Capital cost (kEUR/MW), C^{cc}	Operational lifetime (years), N	Annualized capital cost (kEUR/MW), C^{acc}, at discount rate, $R=0.07$	Fixed cost (kEUR/MW), C^{fix}	Variable cost (EUR/MWh), C^{var}
Nuclear	4661.25	60	332.02	60	18
Coal	1560	40	117.01	42	46.13
Gas CCGT	765.75	30	61.71	20	66.51
Gas COCGT	375	30	30.22	19.5	90.20

Multi-period optimization (dynamic model):

The dynamic model minimizes the total system cost c_{sys}^{total} over several periods. Each period T belongs to the horizon H. For a conventional energy mix and an inelastic demand, the simplest form of the multi-period optimization can be represented as follows:

$$\min_{i_{g,T}^{cap},p_{g,b,T}} \quad c_{sys}^{total} = \sum_T \frac{1}{(1+R)^T} \sum_g \left(\left(\left(1-\frac{T}{N_g}\right) C_{g,T}^{cc} + C_{g,T}^{fix} \right) i_{g,T}^{cap} + \sum_b C_{g,T}^{var} p_{g,b,T} \cdot d_b \right) \tag{2.3a}$$

$$\text{s.t.} \quad 0 \le p_{g,b,T} \le i_{g,T}^{cap} \tag{2.3b}$$

$$i_{g,T}^{cap} = \sum_{j=0}^{T-1} i_{g,j}^{cap} = \sum_{j=0}^{T-1} e_{g,j} P_{g,j}^{max} \tag{2.3c}$$

$$\sum_g p_{g,b,T} = P_{b,T}^d \tag{2.3d}$$

$$b \subseteq T \tag{2.3e}$$

$$T \subseteq H. \tag{2.3f}$$

In contrast to the static model, the multi-period optimization is able to incorporate the changes in C^{cc} (e.g. due to technological improvement) and represent their present value by using the term $1/(1+R)^T$. Moreover, the dynamic model accounts for asset depreciation (i.e. in this case, straight-line depreciation, T/N_g, where it has been assumed that the operational lifetime is equivalent to the financial lifetime of the asset) within the horizon H. In this way, technologies are compared based on how costly they are and how much value is left from the technology at the end of the horizon (known as *salvage value*). With regards to the mathematical formulation, Eq. (2.1c) in the static model has been replaced by Eq. (2.3c) in order to account for the inter-temporal relation between the installed capacities. This means that all power plants built prior to T still exist at T (see that the formulation above does not consider decommissioning or end of operational life) and therefore can be considered for operation.

2.2.4. ESOMs: Technical limitations and potential solutions

As indicated in Subsection 2.2.1, ESOMs and energy system models in general are typically used to represent several regions, technologies and sectors: i.e. some applications covered entire continental areas (e.g. Africa and Latin America) as well as large countries (e.g. China and the U.S.A.), sector coupling and extended horizons (e.g. up to 2050). Since the required computational resources to perform these studies are significant, these models therefore tend to make use of low-time resolution foresights and formulations that explicitly avoid short-term operations equations [29].

The influence of these two limitations has been investigated and compared in the context of high renewable energy integration. For instance, Ref. [68] found that the lack of specific short-term operations is more influential on the ESOM's results at low levels of integration. On the other hand, the influence of the low-time resolution becomes more evident the higher the integration. To further elaborate on this aspect, the two characteristics are subsequently discussed in more detail, together with the limitations they represent and the potential solutions that have been proposed to circumvent them.

Low-time resolution:

To explain the concept of low-time resolution, the load duration curve (LDC) introduced in Subsection 2.2.3 is used as an example. An LDC usually represents one year of power consumption data at one-hour resolution. This means that the LDC can be interpreted as consecutive energy blocks in which each block has a one-hour base and a height equivalent to its corresponding power data point. As aforementioned, however, this high-time resolution becomes computationally demanding for the typical uses given to ESOMs. Hence, energy blocks are built instead by selecting a fixed power data point and extending the duration of its base for more hours than just a single one so that this new block, called *timeslice* in the literature, represents several data points at once. Fig. 2.3 offers a visual comparison between a *timeslice* and the original energy block.

Selecting the number of timeslices, their height (what power data point) and the duration of their bases can be done in several different ways. Prior to the increasing installation of renewable energy sources, the chosen number of timeslices for energy system studies were a few: 1, 6 or 12 [4]; while their power data points were often specific load conditions such as weekend nights in winter [69]. With an increasing deployment of renewable sources, the representation of the LDC itself does not pose any additional challenges, but its representation in relation to the time-dependent behavior of these intermittent sources does, as explained next.

Investment in renewable energy sources can be considered in ESOM's in two ways: 1) by assuming a defined installed capacity so that its influence can be directly injected in the LDC (this refers to the net load

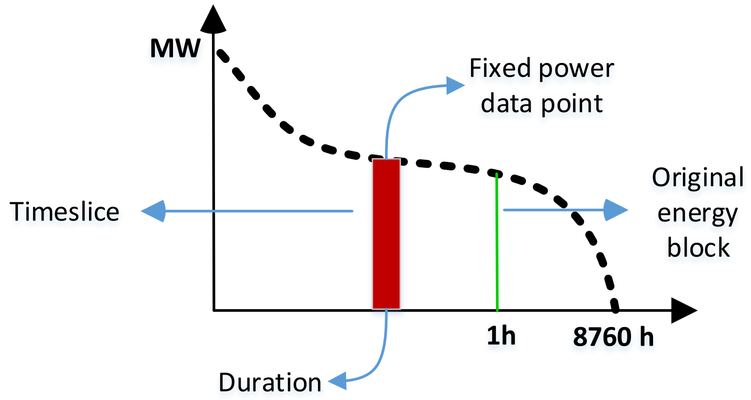


Figure 2.3: Load duration curve (LDC), in which the *timeslice* concept is introduced in comparison to the original energy block. A timeslice is a single power data point with an extended duration, used to lower the sampling resolution of the curve.

duration curve, NLDC, mentioned in Subsection 2.2.3); and 2) by letting these sources compete with other technologies in the optimization problem, while their *resource availability curves* (e.g. their maximum power at each timeslice) are specified. Nonetheless, since neither an accurate temporal representation nor profile variations can be reasonably approximated with a few slices, new studies in the field have increased the time resolution up to 260 [70] or even 288 [71] timeslices per year .

Although the increasing number of timeslices significantly improves the time resolution at the expense of more computational requirements, the best way to capture the relation between the load and variations in renewable energy sources remains an open research question. Some efforts to tackle this challenge have used heuristics [72], optimization [73], machine learning [74] and data clustering [75]. The method selected in this thesis is described in detail hereinafter in Section 4.2, as an introduction to the simulations.

Lack of short-term operations equations:

As illustrated in Subsection 2.2.3, equations representing the single-period and multi-period central generation expansion do not incorporate detailed constraints regarding the operation of power plants or grid power flows. Historically, results obtained with such reduced level of operational details were often deemed sufficient for policy development and even capacity expansion planning in some cases [4]. For instance, short-term flexibility adequacy was fairly omitted in terms of security of supply until recently, as the need for flexibility was mainly driven by power plant and transmission network outages and the variations in demand [6]. In this regard, systems with certain types of power plants (e.g. dispatchable hydro power) may have been considered to posses an inherent capability to meet most power reserve requirements, despite the lack of explicit flexibility considerations within the ESOMs [4]. However, the rapid deployment of renewable power sources and their projected substantial increase in installed capacity [76] can quickly render most of these assumptions obsolete, particularly because these sources negatively affect flexibility in two ways that reinforce each other [6]: 1) they increase the need for system flexibility in order to cope with their sudden output changes and 2) they decrease the conventional supply of flexibility, as they displace power plants that previously provided the service.

As aforementioned, ESOMs have shown to often inadequately represent the need for flexibility [29], but also reserve capacity requirements and inflated costs of ancillary services in power systems with high levels of renewable energy [77], which has effectively resulted in sub-optimal power plant dispatch, underestimated costs of integration and consequently, sub-optimal investment decisions [28, 29, 77]. Furthermore, the lack of more detailed equations does not allow to grasp the impact that net load variations have on power plants, such as reduced efficiency and operational lifetime due to frequent cycling [77]. In this context, they also fail to capture financial implications of inter-temporal operating decisions such as power plant start-ups [78] and steep and often ramping actions [6], as well as the increased maintenance costs that this unusual operational behavior may entail [78]. The misrepresentation of these aspects in energy-only markets (e.g. as described in Subsection 2.2.3), where power plants bid based on their marginal cost, may send improper market signals to power plants investors, whose return of investment has been already significantly aggravated in recent years due to plummeted capacity factors in systems with decreased net loads [79, 80].

A potential solution to incorporate the operational information described above is to enhance these models via heuristic approaches that include short-term operations in long-term simulations (e.g. see Ref. [4]) while still neglecting several details. Another solution is to enrich the outcomes obtained by ESOMs with power system models, which contain more detailed equations that are capable of explicitly calculating short-term metrics, maintenance outages, forced outages rates, mean time to repair modes, load information, ancillary services, system reserve requirements, power plant technical constraints (e.g. minimum up/down time, maximum and minimum power output, ramping constraints), etc. [81] and even other sources of flexibility such as storage [61] and demand response [82].

2.3. Power system models

As described in Subsection 2.2.4, the performance of ESOMs is limited by its low-time resolution and the lack of short-term operations equations. While the low-time resolution can be enhanced by choosing appropriate timeslices, the most accurate representation of the load and renewable sources profiles can in fact only be achieved at full resolution, i.e. 8760 hours. Furthermore, while the missing equations in the model can be approximated by heuristics, the truthful operational representation can only be obtained by specific, usually interlinked detailed equations. In this context, power system models play a fundamental role due their ability to capture these two aspects and may thus be exploited to obtain finer results in conjuction with the ESOMs. Here, the specific unit commitment and economic dispatch (UCED) model will be discussed, as their short-term operations are the most influential on long-term investments on the power sector [9].

2.3.1. UCED: Unit commitment and economic dispatch

The unit commitment problem consists in obtaining the optimal schedule of generators that are to supply the ineslatic power demand over a given horizon. The problem is typically solved by a central algorithm that minimizes a set of costs (e.g. start-up costs, shut-down costs, variable costs, emission costs, etc.), the amount of shedded load or the levels of curtailed renewable power; while active constraints such as ramping rates, minimum up time, minimum down time, mininum power output, maximum power output and spinning reserves are ensured [83].

When the network considered is a copper plate (i.e. no grid constraints), the problem is typically called *unit commitment* (UC). On the other hand, when grid power flow limits and voltage violations are taken into account, the problem is a *security-constrained unit commitment* (SCUC) [84]. If only active power is of importance, both problems can be solved via DC approximations. However, if voltage violations are also to be included in the SCUC, a full AC formulation is required. The simpler UC may be referred as a *market simulation* [85] in reference to how upper level power trading neglects grid constraints in electricity markets. In this context, the SCUC then refers to the optimization that a central operator (e.g. transmission system operator, independent system operator) implements to guarantee system reliability, stability and security after the market has settled.

The unit commitment problem may also be solved in a distributed manner (e.g. Ref. [86]) or from another perspective than that of a central operator. In this latter regard, a particularly interesting concept is the *price-based* unit-commitment (PBUC), which refers to the power plant scheduling problem that a generating company (GENCO) solves in order to dispatch their power assets in response to a price signal generated by the electricity market. In this formulation, the GENCO is no longer obliged to supply hourly demand power levels and instead its objective becomes a profit maximizing function [87].

In this thesis, the classic *unit commitment* is adopted. This approach is widely discussed in the literature and may be solved by deterministic methods (e.g. Priority list, Dynamic programming, Lagrangian relaxation, Integer programming) or heuristic methods (e.g. Tabu search, Simulated annealing, Fuzzy systems, Genetic algorithms, Evolutionary programming, etc.) [84, 88]. Herein, the mixed-integer linear programming formulation is briefly presented [89]:

$$
\begin{aligned}
\min_{p_{g,b},s^{up}_{g,b}} \quad & \sum_g \sum_b \left(C^{var}_g p_{g,b} \cdot d_b + C^{sup}_g I_g s^{up}_{g,b} \right) && \text{(2.4a)}\\
\text{s.t.} \quad & u_{g,b} P^{min}_g \leq p_{g,b} \leq u_{g,b} P^{max}_g && \text{(2.4b)}\\
& s^{up}_{g,b} = \frac{1}{2} \left(\left(u_{g,b} - u_{g,b-1} \right) + \left(u_{g,b} - u_{g,b-1} \right)^2 \right) && \text{(2.4c)}\\
& p_{g,b} - p_{g,b-1} \leq P^{up}_g && \text{(2.4d)}\\
& p_{g,b-1} - p_{g,b} \leq P^{down}_g && \text{(2.4e)}\\
& d_{g,minUp} \left(u_{g,b} - u_{g,b-1} \right) \leq \sum_{b}^{d_b + d_{g,minUp}} u_{g,b} && \text{(2.4f)}\\
& d_{g,minDown} \left(u_{g,b-1} - u_{g,b} \right) \leq \sum_{b}^{d_b + d_{g,minDown}} \left(1 - u_{g,b} \right) && \text{(2.4g)}\\
& \sum_g p_{g,b} = P^d_b && \text{(2.4h)}\\
& b \subseteq T && \text{(2.4i)}\\
& T \subseteq H && \text{(2.4j)}
\end{aligned}
$$

Eq. (2.4b) enforces lower and upper active power limits when the power plant is operating. Eq. (2.4c) obtains the start-up binary variable $s^{up}_{g,b}$ in terms of the commitment variable $u_{g,b}$. Eq. (2.4d) and Eq. (2.4e) restrict the inter-temporal increase or decrease of the active power output (i.e. ramping). Eq. (2.4f) entails that a power plant must remain online for a minimum amount of time if it is to be operational, whereas Eq. (2.4g) implies that the plant must remain off for a minimum amount of time if it is not to be committed. Finally, Eq. (2.4h) ensures that the inelastic power demand is supplied at every b.

Due to the mixed-integer nature of the unit commitment, the search for an optimal solution quickly runs into a combinatorial explosion due to several generating units in the system and multiple time steps (e.g. 8760 hours). To circumvent the complexity regarding the time steps, the optimization is approached using a rolling horizon [90], which consists in breaking the total number of time steps into smaller groups called *horizons* so that the latter are individually optimized at a much lower computational effort (e.g. 8760 hours broken down in 365 horizons of 24 hours each). The horizons can also be overlapped (much lower than the length of the horizon, e.g. 2 hours) to increase the accuracy of the solution. Ref. [85] states that if the length of the horizon is large enough so that minimum down time constraints of generators are within, the solution provided by the rolling horizon method will be close, if not the same, to the actual solution of the problem when solved at once.

To further reduce the problem state-space dimensionality, similar generating units in a copper plate grid can be grouped in clusters [28] while respecting each generator's individual constraint [85]. In this way, the optimization consists of a number of integer variables per cluster that is much smaller than the number of binary decisions per generator. Ref. [28] calculates that the reduction in the decision space is considerable: from the total number of binary decisions per generator 2^{num_gen} to the product of the cluster sizes

$\Pi_{clust} cluster\_size_{clust}$. For example, Ref. [91] applies a similar reduction to the Korean power system: 140 power plants in fewer than 10 clusters. Ref. [91] argues that such approach is acceptable since manufacturers produce power plants with pre-established sizes, types and seemingly similar constraints so they can be easily grouped.

2.4. Computational tools

2.4.1. ESOM: OSeMOSYS

The energy system optimization model selected for this thesis is the Open Source Energy Modelling System (**OSeMOSYS**), introduced in 2011 in Ref. [92]. OSeMOSYS is chosen because its code is flexible, easily understandable, it is freely distributed (and equally important, written in a free-of-charge software) [2]. OSeMOSYS's initiative is currently supported by a dedicated community (`groups.google.com/forum/#!forum/osemosys`) and an official website (`http://www.osemosys.org/`). The software is a mature long-term energy optimization model that has been used for several energy planning studies (see Refs. [30, 35, 36, 94–96]).

The structure of OSeMOSYS consists of **INPUTS**, **MODEL** and **OUTPUTS**. The inputs can be further divided in **SETS** and **PARAMETERS**. **SETS** are labels assigned to the model such as the name of the region, types of pollutants (e.g. CO_2, NO_x, etc.), fuels (diesel, oil, electricity, etc.) and technologies (coal power plant, gas power plant, gas extraction, gas import, etc.), which remain constant throughout the optimization period. On the other hand, **PARAMETERS** include capital, fixed and variable costs, emission penalties, renewable production minimum targets, capacity factors, availability factors, etc., which may have different values during the model periods.

MODEL is composed of constraints and a single objective function. The constraints ensure that the energy system model has sufficient capacity adequacy and energy adequacy and that parameters inserted by the modeler (e.g. total installed capacity of a technology) are respected by the problem. The objective function is the total discounted cost of the whole system. The total discounted cost in OSeMOSYS is the difference between all discounted costs (capital costs, operating costs and emission penalties) and the discounted salvage value of technologies and storage at the end of the modeling period. Finally, after OSeMOSYS has solved the optimization problem, it provides several outputs for further processing such as the value of the objective function, the new capacity installed and the energy produced per technology, the investments per sector, infrastructure, technologies, storage and most interestingly, when to invest on them.

OSeMOSYS was originally written in the GNU MathProg language and subsequently translated to GAMS and Python. According to OSeMOSYS developers in the Q&A community, the GNU version is the official one and therefore, maintenance and approved improvements to the software may only be applied to that version. In this thesis, the first tests regarding OSeMOSYS were implemented using GNU MathProg. In this version, OSeMOSYS and the input data are both **.txt** files, which are matched together with a solver via

[2] In contrast, it is possible to find similar ESOMs such as MARKAL/TIMES whose model has a steeper learning curve. Furthermore, although its code is free of charge too, the software where it is written is not [93].

the command window, **cmd**. It was soon realized that this method has an significant limitation: parameters and variables cannot be easily changed externally due to the nature of **.txt** files. Therefore, the GNU version was not considered to be the best choice if more dynamic simulations with multiple interactions were needed. For this reason, the following tests were performed using the Python-Pyomo version, which requires a **.dat** file as an input that is similarly not so easily modifiable.

In an effort to circumvent the problems that Python-Pyomo and GNU posed, OSeMOSYS-PuLP was found in Ref. [97] and `https://github.com/OSeMOSYS/OSeMOSYS_PuLP`. This version is entirely written in Python-PuLP and includes the option of performing Monte-Carlo simulations. In this version, the parameters and sets are inserted in **.csv** files (very user friendly) and are loaded by the code to the Python variable space as dictionaries. Thus, the problem parameters, sets, variables, constraints and overall optimization problem are all easily readable and modifiable in the Python work space. This was the final decisive factor for choosing OSeMOSYS-PuLP over all other versions.

2.4.2. PSMs: PowerFactory, PyPSA

PowerFactory 2019:

PowerFactory is a commercial software developed and maintained by the German company, DIgSILENT GmbH. The software offers several functions such as load-flow and short-circuit analysis, state estimation, motor starting, small signal stability, electromagnetic transients, etc. In their 2019 release, the advanced functionality **Unit Commitment and Dispatch Optimisation** was launched. This newly added tool opened up the possibility to further explore this software and enhance the author's previous experience in it. Furthermore, PowerFactory is commonly used in the Intelligent Electrical Power Grids (IEPG) group at the Delft University of Technology, where the author is currently doing his research. This aspect was considered essential since the accumulated knowledge in the group could be beneficial when in doubt. Taking these factors into account together with the strong technical rigorosity and vast amount of short-term operations details provided by PowerFactory, it was decided that this software is a particularly good choice for the research here implemented.

The unit commitment toolbox offers two types of single objective functions to be minimized: user-defined and total costs. The *user-defined* does not allow to insert an actual objective function of preference as its name suggests, but choose instead from a list of single-objective functions such as the minimization of generator operating costs, generator start-up/shut-down costs, minimization of renewable power curtailment, etc. On the other hand, the *total costs* intends to minimize all possible costs in the system, which may include variable costs, start-up costs, shutdown costs, cost of load shedding, cost of renewable power curtailment and cost of re-dispatch (for example, when considering the intra-day market). The search space of the objective functions can be constrained by generators' physical limits (e.g. minimum and maximum active and reactive power dispatch, inter-temporal ramping, start-up and shut-down limit), as well as by grid restrictions such as exceeding line flows, voltage violations, etc.

Moreover, the unit commitment toolbox allows to solve for AC or DC unit commitment. For market simulation, it is advised to use the DC formulation and subsequently assess the grid's influence on the power dispatch via an AC calculation. The optimization problem can be solved by using incorporated solvers such as the open-source **Cbc** and **lp_solve**, but PowerFactory also provides an interface to use external ones, especially for commercial, but much more powerful solvers such as CPLEX and Gurobi. In this thesis, it was realized that using Cbc, the simulation time to solve the unit commitment problem was approximately three times faster than in other simulation tools (see e.g. PyPSA, explained below), which makes PowerFactory very attractive when real-size power systems are considered.

However, as per the 2019 release, the unit commitment problem is only focused on generators and the grid, leaving aside the possibility to include storage units. Storage was only introduced in the subsequent release, PowerFactory 2020. Given that the possibility to *easily* (e.g. not considering complex ways around it such as creating a dynamic model with differential equations) include storage in the simulations was a limiting factor to analyze different effects on the system, the author decided to explore other options such as the open-source software Python for Power System Analysis (PyPSA).

Python for Power System Analysis (PyPSA):

Python for Power System Analysis (PyPSA) is an open-source software, currently maintained by the Energy System Modelling group at the Institute for Automation and Applied Informatics at the Karlsruhe Institute of Technology in Germany. The toolbox has a main website (`https://pypsa.readthedocs.io/en/latest/index.html`) where its functionalities and attributes are duly explained. When in doubt, it is possible to pose questions to the developing team at the following specific forum: `https://groups.google.com/forum/#!forum/pypsa.` Beside these extremely helpful means to independently gain experience in PyPSA, this software is also selected because it offers understandable models that are Python-based, and in particular, it allows to easily incorporate storage units in the simulation (i.e. the very reason why PowerFactory 2019 is to be replaced for this part of the thesis).

PyPSA has a single optimization function that minimizes the total system costs. These costs vary depending on whether the toolbox is being used as an investment optimization or as a power system operations optimization. For example, if the tool is to be used for investment, the total system costs include the capital and variable costs of generators, storage and transmission; while the software provides the optimal mix as a result. However, if the tool is to be used for operations optimization only, the technology sizes are provided by the modeler; thus, the capital costs are disregarded, but the variable costs remain as a decisive factor. PyPSA also offers an option for unit commitment, but this can only be used for power system operations and its use for investment decisions is strictly forbidden as this couples decision variables and linearity is lost. A way around this could be to use the clustered generation expansion planning, proposed by Ref. [28].

In PyPSA it is possible to optimize as many snapshots (in this thesis, this is equivalent to t) as needed, but as of version 0.17.0, PyPSA does not account for the time value of money or asset depreciation, so it cannot be *naturally* used for a full optimization horizon (several years), as opposed to OSeMOSYS. However, a very useful trait about PyPSA is that it offers the possibility to modify the objective function or write a completely new objective at will, as well as adding specific constraints (e.g. CO_2 emissions).

PyPSA provides models for generators, transmission lines, loads, transformers, buses and storage units. It also offers two fundamental components: links and stores. Links can be interpreted as edges in a graph that can carry any sort of carrier. Their power capacity is an optimization variable during investment decisions. Stores can be interpreted as sinks, which store energy from a given carrier. Their energy capacity is an optimization variable during investment decisions. With these two components it is fairly simple to model any sort of storage device. The technical limitations of each of these models set the constraints for the objective function.

PyPSA is capable of modeling other energy sectors, but in general, it is moslty oriented towards the electricity sector, so it is not *strictly* an energy system model. On the other hand, its electric models and equations are far from the details provided by PowerFactory, so it cannot be fully categorized as a power system model either. In fact, PyPSA is a hybrid that was created to bridge the gap between power system analysis software and energy system modelling tools [89]. Thus, it purposely avoids technical overcomplications (e.g. for instance, it uses linear power flow equations only), but it keeps the essence of the problem in such a way that it remains relevant for longer-term decisions.

PyPSA uses Pyomo to set the optimization problem, while it is able to use any solver that Pyomo understands (e.g. CPLEX, Cbc, GLPK, Gurobi) to find the solution. The developing team has realized that the use of Pyomo is time consuming and have proposed new formulations to access the solver directly. A specific example of some of the time complexity difficulties that PyPSA faces is the unit commitment problem, which cannot bypass Pyomo in the 0.17.0 version. While comparing PyPSA with PowerFactory, using the same conditions (e.g. solver, rolling horizon, overlap, number of generators, etc.), the solving procedure roughly lasted three times more than it took Python to access PowerFactory, solve the problem in PowerFactory's engine and retrieve the results back to the Python's workspace. The author tried several times to improve this by parallelizing the unit commitment computations, but the task proved to be much more highly complicated than expected and not to be necessarily related to the purpose of this thesis.

2.4.3. Comparison overview

Subsection 2.4.2 provided a detailed explanation of the structure of the computational tools used in this thesis, as well as some of the opinions and experiences from the author regarding their use. Following this same methodology, Table 2.4 provides a comparison overview of the three computational tools in order to help the reader better assess their advantages, limitations and differences.

Table 2.4: Comparison overview between OSeMOSYS-PuLp, PowerFactory 2019 and PyPSA 0.17.0. While many of these considerations are factual, some do represent the views of the author and may not necessarily hold true for other applications. 'N/A': not applicable, '+': acceptable/possible, '++': good.

Aspect	**Specific**	**OSe-PuLp**	**PF 2019**	**PyPSA 0.17.0**	**Comments**
Investment decisions	Multi time-step optimization	++	N/A	++	This refers to multiple time snapshots, e.g. hours in a single year. OSeMOSYS and PyPSA can easily perform such optimization.
	Multi period optimization	++	N/A	+	This refers to several periods in a single horizon, which allows to consider the time value of money and asset depreciation. OSeMOSYS naturally incorporates this aspect, whereas PyPSA could do it by externally adding it.
	Other energy sectors	++	N/A	+	OSeMOSYS is sector agnostic, whereas PyPSA allows to model other sectors to a limited extent through their fundamental components: links and stores.
Power system operations	Detailed models and equations	N/A	++	+	Models in PowerFactory are of extremely high quality. Nonetheless, such level of detail is not precisely needed for investment decisions. PyPSA purposely omits these overcomplications.
	Storage	+	+	+	PowerFactory does not have a built-in storage component for investment decisions. OSeMOSYS and PyPSA model storage very similarly, but the latter allows for more parameters to be included, such as the standing loss.
Flexibility with respect to the user's needs	Objective function and constraints	+	+	++	PowerFactory has built-in objective functions and constraints that cannot be changed. It is not flexible per se, but it offers multiple options for optimization. On the other hand, the open source nature of OSeMOSYS and PyPSA allows to easily modify the code at will. However, PyPSA in fact, offers an enhanced functionality that allows to directly modify both the objective and constraints without necessarily accessing the base code.

Unit commitment (UC)	Speed	N/A	++	+	Under the same conditions, the speed attained by PowerFactory to solve the UC problem is about 3 times faster than that of PyPSA. A possible explanation for this is the use of Pyomo by PyPSA and its non-parallel computations, which significantly hinder the software's speed in this regard.
	Formulation	N/A	++	+	Both PowerFactory and PyPSA are capable of solving security-constrained UCED. However, PowerFactory offers the possibility to solve a full AC problem (therefore, obtaining voltages and reactive power), whereas PyPSA only solves linear power flow equations.
	Clustering	N/A	++	N/A	PowerFactory allows to cluster generators in virtual power plants. In this way, a whole fleet of generators can be considered for commitment instead of single machines, which significantly reduces the computation time. PyPSA does not offer this option.

3
Models: Coupling

3.1. Introduction

In Chapter 2, the energy system optimization model (ESOM) and the unit commitment and economic dispatch (UCED) model were introduced. As it was described, ESOMs allow to model long-term investment decisions for several energy sectors, but have two main limitations due to their typical uses: low-time resolution and lack of short-term operations equations. On the other hand, UCED models do facilitate the assessment of short-term power system operations, including flexibility requirements and detailed costs of operation. Therefore, UCED models could be used together with ESOMs to help understand the influence of short-term operations on long-term investments, as first posed in Chapter 1.

Within this context, this chapter thus provides a more specific definition about this coupling scheme between models; it elaborates on a family of methods proposed by other scientific works to implement such coupling; it discusses about other potential methods with different characteristics; it advocates for the selection of one of the methods; and finally, it explains how this preferred method is employed with the chosen computational tools (OSeMOSYS, PowerFactory and PyPSA).

3.2. Coupling methods: Concept and review

The coupling scheme here adopted makes three assumptions: 1) the models considered are solely ESOMs and UCED models, 2) at least one of the models is influenced by the outcome of its counterpart and 3) a model structure may be modified only prior to the information exchange. The first assumption is to focus on a specific niche of models and create defined boundaries between other model combinations that may seem very similar: e.g. Ref. [98, 99] couple an energy system *simulation* model with an UCED model. The second and third assumptions are to further narrow the research scope. Nonetheless, some works that do not follow this latter condition are still presented in *Other forms of coupling* (Subsection 3.2.4) to offer a ground for comparison.

The coupling scheme consists of three different categories, compiled in Ref. [100] [1]: (1) co-optimization, (2) undirectional softlinking and (3) bidirectional softlinking. Fig. 3.1 offers a graphical representation of the concepts behind these coupling methods. As observed, all parameters and variables from the ESOM and UCED models constitute a single optimization space in the co-optimization approach. On the other hand,

[1] In fact, Ref. [100] includes several more divisions, but some are not particularly related to coupling and others do not adhere to the two assumptions made here.

variables from the ESOM are passed to the UCED model as parameters in the unidirectional soflinking method, while the bidirectional softlinking mechanism performs a similar forward coupling between the ESOM and the UCED model, but it closes the loop by also adding a feedback flow of information. That is, this feedback loop passes variable results from the UCED model to the ESOM parameters. To further elaborate on these methods, the challenges encountered and the ideas proposed in different scientific works are discussed next.

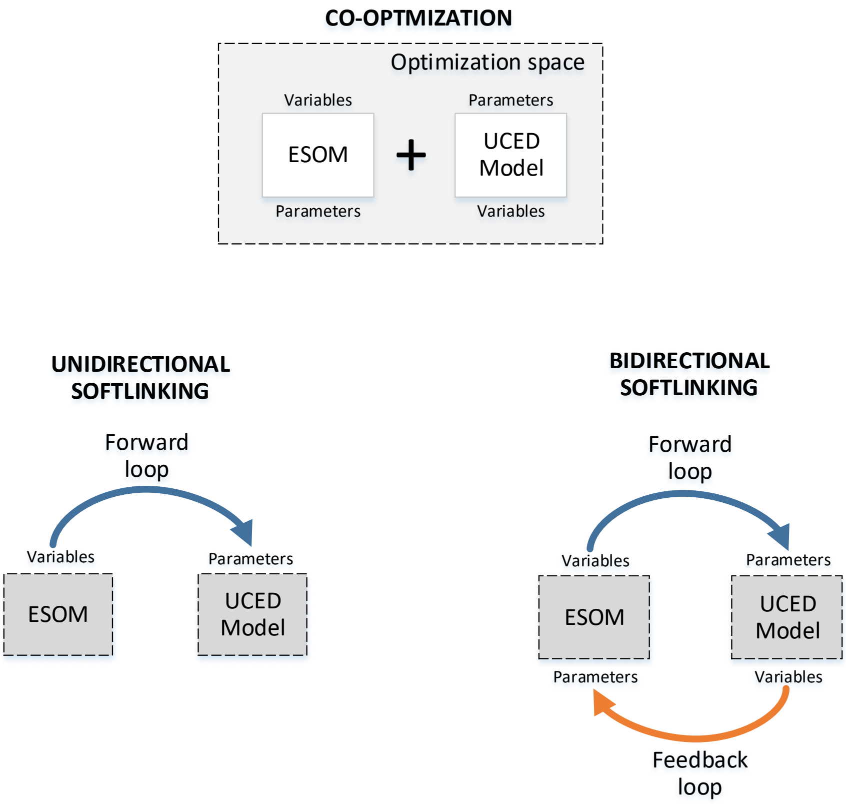


Figure 3.1: Coupling methods between an energy system optimization model (ESOM) and a unit commitment and economic dispatch (UCED) model. The co-optimization approach solves a single optimization problem. On the other hand, the unidirectional softlinking forwards the variable results from the ESOM to be inserted as parameters in the UCED model, while the bidirectional softlinking closes the loop by adding a connecting link in the opposite direction (i.e. feedback link).

3.2.1. Co-optimization

As shown in Fig. 3.1, the co-optimization coupling method merges the formulation of ESOMs and UCED models in a single optimization space. For instance, consider the combined mathematical formulation of the ESOM sub-problem, i.e. the single-period central generation expansion planning (CGEP) introduced in Subsection 2.2.3, and the UCED model described in Subsection 2.3.1.

$$\min_{i_g^{cap}, p_{g,b}, s_{g,b}^{up}} \quad c_{sys}^{total} = \sum_g \left(\left(C_g^{acc} + C_g^{fix} \right) i_g^{cap} + \sum_b \left(C_g^{var} p_{g,b} \cdot d_b + C_g^{sup} i_g^{cap} s_{g,b}^{up} \right) \right) \tag{3.1a}$$

$$\text{s.t.} \quad \begin{matrix} Eq.\ (2.1b), Eq.\ (2.1c), Eq.\ (2.1d), Eq.\ (2.1e), Eq.\ (2.1f) \\ Eq.\ (2.4b), Eq.\ (2.4c), Eq.\ (2.4d), Eq.\ (2.4e), Eq.\ (2.4f), Eq.\ (2.4g) \end{matrix} \tag{3.1b}$$

$$u_{g,b} \leq e_g \tag{3.1c}$$

The newly added constraint, Eq. (3.1c), connects both problems since it ensures that a generator can be committed ($u_{g,b} = 1$) if and only if it is has been built ($e_g = 1$). This type of coupling is able to maintain linearity by strictly requiring installed capacities (i.e. Eq. (2.1c)) to be discrete. If this latter condition was not to hold, the continuous i_g^{cap} would replace P_{max} in Eq. (2.4b), thus linking both problems by a non-linear constraint.

The combined formulation of both models is inclined to quickly run into a combinatorial explosion faster than the original UCED model due to the additional binary variable (i.e. e_g) per timestep. To bypass this obstacle, Ref. [28] proposes a clustered generation expansion planning, which makes use of a clustered unit commitment as described in Subsection 2.3.1 and a clustered power plant maintenance schedule. The flexibility provided by thermal power plants in the context of high integration of renewable energy sources is one of the main points of discussion in this work. A similar approach can be found in Ref. [6], where other sources of flexibility are also included: supply-side (dispatchable conventional power plants and curtailable renewable sources), demand-side (prosumer and consumer demand response), energy storage and power systems interconnection.

As in Eq. (3.1), it is similarly possible to couple the multi-period CGEP described in Eq. (2.3) and the UCED model, with a natural increase in complexity. An example of this coupling is found in Ref. [69], which presents an integrated problem to minimize the total discounted cost (variable, emissions, import costs, etc.) for the Greek power system, including its interconnections with neighboring countries. To lower the computational complexity, the grid is split in two sectors and five zones, while each month in a year is represented by one 24h day. A comparable study is found in Ref. [61], in which the optimization equally searches for a minimum total discounted system cost. However, this formulation is more technologically populated in contrast to Ref. [69], as it also considers pumped-hydro and battery storage. In this case, the problem complexity is reduced by using representative intervals for the time resolution and a linear version of the UCED model.

Another type of co-optimization is found in augmented versions of ESOMs that try to replicate themselves some of the results obtained with the UCED models. For instance, Ref. [77] uses the new dispatch and unit commitment extension of the bottom-up energy system optimization model, Integrated MARKAL-EFOM System (TIMES). This built-in feature allows to select between continuous binary commitment variables or fully discrete. In the discrete version, the traditional short-term constraints like start-up time, ramp rates, minimum load level, shutdown time, minimum online and offline time are included [77]. Another example in this realm is found in Ref. [4], where OSeMOSYS is enhanced to include short-term dynamics using heuristics. Reliability and reserve-requirement assessments are implemented. The former is incorporated via an analytic approximation of wind capacity credit, whereas the latter is determined based on forecast errors from supply and demand that are assumed to be normal distributed. This procedure allows to consider the maximum upward and downward primary and secondary reserve contributions of each technology as well as their minimum stable generation levels [4].

3.2.2. Unidirectional softlinking

A different approach from the one proposed by the co-optimization coupling method is to keep the ESOM and the UCED model separate, each with their corresponding characteristics. A link can thus be established from the ESOM to the UCED model, in which variable results calculated by the former are passed as parameters to the latter. This coupling method is known as *unidirectional softlinking*. A specific example of this forward link in the electricity sector is the installed capacity per technology calculated by the ESOM, which is transferred to the UCED model as a parameter for further assessment: e.g. capacity adequacy, reliability and actual operating costs (while taking into account all operational constraints).

A benchmark in this coupling method is Ref. [81], which unidirectionally softlinks the Irish energy system modeled in TIMES and the Irish power system modeled in PLEXOS. This study quantifies the effects of high integration levels of wind on reliability (security and supply adequacy) and flexibility requirements on the energy portfolio. Their work is based on a seven-step methodology, which is summarized as follows: pick a specific year from the ESOM solution to be analyzed in detail by the UCED model; run the UCED model from a no-constraints case to a full-constraints case; and finally, evaluate the influence of multiple and different (usually low) wind profiles on the energy mix.

The same methodology is applied to the Italian power sector in Ref. [79], using TIMES and PLEXOS alike. However, this study is more comprehensive, because it also includes many other energy sectors from the Italian economy in order to assess their effects on the development of the power system and vice versa. It is then declared that the end goal is to measure potential alternative strategies that eventually lead to an improved performance of the power system, which is measured via reliability metrics, electricity prices, flexibility requirements, etc.

In this same line of methodology, Ref. [80] unidirectionally softlinks MARKAL-NL-UU (MARKAL model specifically developed for the Dutch energy sector; see Ref. [101]) and REPOWERS, a UCED model. Their focus lies on the operation of power plants rather than on the reliability of the power system [80]. Four cases are discussed and the results from the UCED model are assessed in a post-analysis. The post-analysis considers four performance metrics: 1) the adequacy of hourly and sub-hourly reserves, 2) the efficiency reduction of thermal power plants, calculation of the 3) annual short-run profit and the 4) discounted payback time for all power plants.

Two more examples of unidirectional softlinking are found in Ref. [68] and Ref. [76]. In Ref. [68], TIMES and

LUSYM (see [102]; an earlier version of LUSYM Invest, aforementioned in the co-optimization coupling method) are linked to study the influence of added operational constraints in the power system in constrast to increasing the time resolution on the ESOM side. On the other hand, Ref. [76] studies the effects of increasing levels of wind and solar integration on the Australian National Electricity Market. Here, the MC-ELECT tool (see Ref. [103]; a Monte Carlo-based CGEP) is used to obtain long-term investment decisions by applying financial portfolio analysis techniques to determine an efficient frontier, which contains optimal generation mixes given possible tradeoffs between expected cost and its associated cost uncertainty [78]. The optimal generation portfolio for each level of integration is then scrutinized using short-term operational constraints from PLEXOS.

3.2.3. Bidirectional softlinking

The bidirectional softlinking coupling method complements the ideas and concepts from the unidirectional approach described above by placing a feedback link that connects the UCED model to the ESOM. This closed loop therefore intends to emulate a single optimization problem (i.e. co-optimization coupling method), while keeping both models apart without the need to establish coupling constraints such as Eq. (3.1c). As aforementioned, this feedback link consists in injecting variable results from the UCED model into the ESOM parameters. The way in which these results are transformed or adapted to fit the structure of the ESOMs are hereinafter referred to as *adjustments*.

An example of bidirectional softlinking is found in [104]. In this article, a decision-tree model is proposed for energy planning, which considers the growth of demand and electricity prices as probabilistic events, whereas changes in the generation portfolio are tree decisions with an assigned net present value (NPV) each. This NPV is obtained by using a CGEP-like formulation. Every portfolio is then passed to a UCED model in Matlab that calculates the profit of each power plant and feeds this information back to the tree in order to adjust the NPV of the decision. The end goal is to find the decision that yields the maximum NPV. In this sense, an established convergence criterion does not exist, but the process rather consists in a search within a given space of options.

A different type of study is found in Ref. [71] for the period 2005-2050 in Portugal. TIMES is used as the ESOM and EnergyPLAN as the UCED model. The ESOM provides the installed capacity per year for each technology, while the UCED model minimizes the variable costs of operation and calculates the levels of renewable integration. Specific attention is paid to the latter, which becomes the main criterion for convergence: if 90% of the energy produced by every renewable technology is consumed, the exchange stops for that specific year and the algorithm moves to the next year. If the criterion is not fulfilled for a specific technology for example, the UCED model calculates what capacity would be needed for that technology to maximize its use in the short-term. With this information in hand, it sends an updated parameter (i.e. maximum capacity allowed) to the ESOM in order to readjust the constraints on its installed capacity and energy activity rate. This process is repeated until the whole modeling horizon has been covered.

In this same context, Ref. [91] couples an ESOM, Wien Automatic System Planning Package (WASP), with a self-developed UCED model with clustered unit commitment. WASP performs a multi-period optimization to find the optimal energy mix based on capital costs, etc. and approximations of the operation and maintenance (O&M) costs. Then, each planning year is further analyzed with the UCED model, which includes exact maintenance, repair, operating, start-up and shut-down costs. This higher resolution information is used to update the O&M parameter in the ESOM. The convergence occurs when O&M costs no longer change.

Similarly, Ref. [105] compares the economic value of several flexibility options (e.g. demand response, storage, building more thermal units and transmission lines) in a large power system with a significant amount of reservoir hydro power. The investment decisions are obtained from an ESOM model, Balmorel, whereas the operation of the power system is assessed with WILMAR Joint Market Model (JMM), an UCED model. In this study, the invesment decisions are directly passed to WILMAR, while the information obtained with the UCED model is fed back via an *adder* (similar to a perturbation), which tries to approximate the impact of missing constraints in the ESOM. The correct adder is presumably the one that gives the total system costs when investment costs and actual operational costs are added together [105]. The chosen adder for the simulations is the variable costs of heat pumps and batteries, as the ESOM uses these two technologies more often than the UCED model. The selection of the exact adder becomes a challenge so a sweep over different values is performed. Results show that the system benefits obtained by feeding back information to ESOM increase only slightly when heat pumps are perturbed, but are much more significant in the case of batteries.

Finally, a more recent example of bidirectional softlinking is found in Ref. [70]. To the best of the author's knowledge, this article interestingly seems to be the first one in the context of ESOM-UCED linking to make a clear use of the term *bidirectional coupling* and to include as well a detailed methodological framework. The country under study is Norway, whose electricity sector is hydropower dominated. Hence, the ESOM used is TIMES-Norway (developed in TIMES and which includes several other sectors of the economy), while the UCED model is the EFI's Multi-area Power-market Simulator (EMPS), specifically designed for multi-area power systems (e.g. Continental Europe) simulations with a considerable share of hydropower. Variables calculated by TIMES are passed as parameters to EPMS. These variables comprehend installed capacities, transmission capacities and the electricity demand (this cannot be passed to EMPS beforehand since other electrified energy sectors may increase or decrease the actual electricity consumption perceived by the power system). On the other hand, variables calculated by EMPS are fed back to TIMES-Norway. These variables are the electricity trade prices between Norway and other European countries (taking into account transmission capacity constraints, for example) and the hydropower constraints. While the trading prices can be inserted directly in TIMES, hydropower constraints need to be communicated via availability factors on the technology. Finally, the process converges when the revenue earned by hydropower producers is equivalent (subject to a tolerance) in both models. The revenue in each model is obtained based on their endogenous calculation of their energy prices, which are much more accurate at the resolution of EMPS.

3.2.4. Other forms of coupling

At the start of Section 3.2, it was established that the three assumptions to delimit the definition of *coupling* in this thesis are: 1) the models considered are solely ESOMs and UCED models, 2) at least one of the models is influenced by the outcome of its counterpart and 3) a model structure may be modified only prior to the information exchange. In the literature, it is possible to find ESOMs-UCED couplings that fulfill the second assumption, but that do not employ the third one.

For example, Ref. [106] manually changes both the ESOM and the UCED model on the run. To begin with, the article seeks to minimize the total generating cost and the total amount of CO_2 tonnes ejected to the environment. The resulting energy mix obtained from the multi-objective ESOM is then passed to the UCED model to evaluate whether the mix is flexible enough for the application considered. If not, the mix is deemed unfeasible and 1 GW gas power plant is introduced to both models. The loop repeats until the mix becomes feasible for the required flexibility.

Ref. [107] similarly modifies the structure of the models on the run. It first calculates the installed capacity per unit with an ESOM, which are then transferred to a UCED model that makes use of an enhanced priority list method developed in Ref. [108] to search for the solution. Once in the UCED model, a perturbation algorithm is applied, which consists in: first removing one power plant at a time (the power plant that leads to the maximum cost decrease is thus permanently removed); and subsequently, adding one power plant at a time to apply the same consideration regarding the decrease in cost. Note that the enhance priority list method plays a fundamental role in this article since the proposed brute-force search would be extremely computational demanding if a traditional UCED formulation (i.e. as in Subsection 2.3.1) was applied.

Another example of structural modification is found in Ref. [109], where an iterative process based on Bender's decomposition is proposed. In this study, an ESOM is used together with a UCED model to find the optimal energy mix. The investment problem is similar to the one presented in Subsection 2.2.3, with an additional factor on the objective function that gives an approximation of the UCED model's influence on the ESOM. This factor becomes closer to its true value in every iteration as the set of generators calculated each time is further scrutinized in an actual UCED run that sends information back to impose more constraints on the factor. The increasing number of constraints thus delimits the factor's feasibility region. These constraints, however, represent an active change of the model's structure.

3.2.5. Why bidirectional softlinking?

As covered in Chapter 1, the argument in this thesis has been crafted around the premise that the short-term power system operations captured in UCED models may influence the long-term investment decisions obtained in the ESOMs. This idea hence suggests that the focus should be placed on two coupling methods: 1) co-optimization and 2) bidirectional softlinking. Unidirectional softlinking, on the other hand, may still be used as a first-instance tool to obtain preliminary results, but not as a stand-alone resource due to its lack of influence over the ESOMs decisions.

The question is then whether to choose co-optimization or bidirectional softlinking. Co-optimization first strikes as a computational demanding problem. This can of course be avoided by using linear versions of the UCED model instead. Following this line of argumentation, Ref. [61] shows that allowing binary commitment variables to be continuous does not compromise the accuracy of the solution for the cases considered in their study, since the discrepancy in the total system cost is only 0.52%. Simultaneously, the gain in the simulation speed is significant, because the average computational time is decreased from 2746s to 100s. However, given the amount of parameters and variables of the joint problem, Ref. [61] must also make use of representative intervals. This then raises the question: if representative intervals are being used just as in the case of ESOMs alone, does the co-optimization approach only improves the operational details?

Other two aspects to consider regarding the co-optimization coupling method are its flexibility to be modified and its potential to represent other energy sectors. Regarding its flexibility it is worth wondering whether the addition of a new constraint (for example, in the short-term power system operations) requires a major restructuring of the overall model and whether this extension may trigger different unwanted effects on previously established equations. Furthermore, it is necessary to question whether a single optimization problem is capable of modeling the electricity sector in conjunction with other energy sectors, particularly in the case when short-term operations of those sectors (e.g. gas networks) are to be included. This could undoubtedly make the co-optimization problem quickly become obfuscated.

For the reasons given above regarding the co-optimization coupling method, this thesis consequently advocates for the use of the bidirectional softlinking method. As shown in Fig. 3.1 and discussed in Subsection 3.2.3, this approach allows to keep both the ESOM and UCED model apart so that they can be extended and modified more freely; that is, the flexibility level of each model is preserved. Nonetheless, it is also possible to find several difficulties in this method. For instance, while the variable results from the ESOM sent as parameters to the UCED model (i.e. the forward link) seem to be usually clear, the information which must flow in the opposite direction represents a significant challenge and not an obvious choice [47]. The feedback link therefore commonly becomes a decision of the modeler.

Another challenging factor to be defined in the bidirectional softlinking method is the convergence criterion: e.g. what specific variables to track? Perhaps an equally important question is whether the iterative process will converge at all: even if the threshold is not restrictive, does the problem ever reach a steady state? Furthermore, it is worth assessing whether the coupling approach is tractable (e.g. how many iterations will be needed?) and whether any possible improvements obtained in the results justify the computational resources (e.g. simulation time, memory usage, etc.) to be used.

3.3. Coupling architecture: Bidirectional softlinking (BSL)

As indicated in Subsection 3.2.5, the bidirectional softlinking (BSL) coupling method has been selected due to the flexibility it offers for model modification and extension, as well as the easiness to include the electricity sector in conjunction with other energy sectors. The concept behind the BSL method was generally introduced in Fig. 3.1: i.e. ESOMs and UCED models are able to mutually influence each other by exchanging their corresponding parameters and variable results.

In this context, Fig. 3.2 offers a more specific depiction of the BSL concept in reference to the computational tools (OSeMOSYS, PowerFactory and PyPSA) used in this thesis. As observed, central to the simulation is a Python script that is in charge of implementing the information exchange, storing results and checking for the convergence criterion. This exiting condition for the simulation is usually the convergence of a relevant value to the modeler (e.g. total system costs, including capital investments calculated by the ESOM and actual operating costs obtained with the UCED model). The convergence of this value is specified by a threshold. In cases when the iteration process approaches a steady-state solution, it is enough to simply compare the results from iteration k with iteration $k-1$ against this threshold. However, it may also happen that the problem is not necessarily well behaved and keeps producing different outputs with no steady-state solution at sight. Thus, a possible work-around is to compare the k iteration with the full array of previous solutions (i.e. if the difference between k and a previous solution is smaller than the threshold). This latter implementation also prevents the simulation from oscillating, if that was the case.

Furthermore, each model in Fig. 3.2 is represented by a box with input ports (i.e. parameters) and output ports (i.e. variable results). Parameters are differentiated between those inserted by the user (in red) and those that are to be updated during the simulation run (in blue). Moreover, each box has one incoming link and one outgoing link that represent the flow of information between softwares. Naturally, OSeMOSYS can feed PowerFactory and vice versa, given that they are an ESOM and a UCED model, respectively. However, a more interesting case is PyPSA, which can not only be softlinked with the other two tools but it may also iteratively update itself, switching from one working state to the other. To further elaborate on the content of the boxes representing the tools, the following subsections provide and idea of how these models can be accessed, changed; and lastly, the results of their calculations be retrieved for further processing and simulation. Note that this is thus complementary to Subsection 2.4.1 and Subsection 2.4.2 where an overview of the models' structure, characteristics and techniques was provided.

3.3.1. OSeMOSYS as a functional block

The original code of OSeMOSYS-PuLp is a single, stand-alone script. As shown in Fig. 3.3, the original script first loads the parameters. These parameters come from a **.csv** file inserted by the user, but also their corresponding default values in case nothing else is specified. Afterwards, the original code creates the constraints and the objective function using Pyomo and then solves the optimization problem. Finally, the code prints the results to another **.csv** file, where variable results are organized and can be easily assessed by the modeler. Note that the Monte-Carlo simulations feature, briefly mentioned Subsection 2.4.1, is ignored.

The simulation procedure of OSeMOSYS as a single script is not the most convenient to include the model in an iterative process. Therefore, the code is transformed into a function in order to implement bidirectional softlinking. The OSeMOSYS script is then saved as a module and the function can be easily accessed by external scripts. As depicted in Fig. 3.3, the new script similarly loads the parameters defined by the user, which are needed to calculate the first ESOM's solution. Nonetheless, the optimization problem is

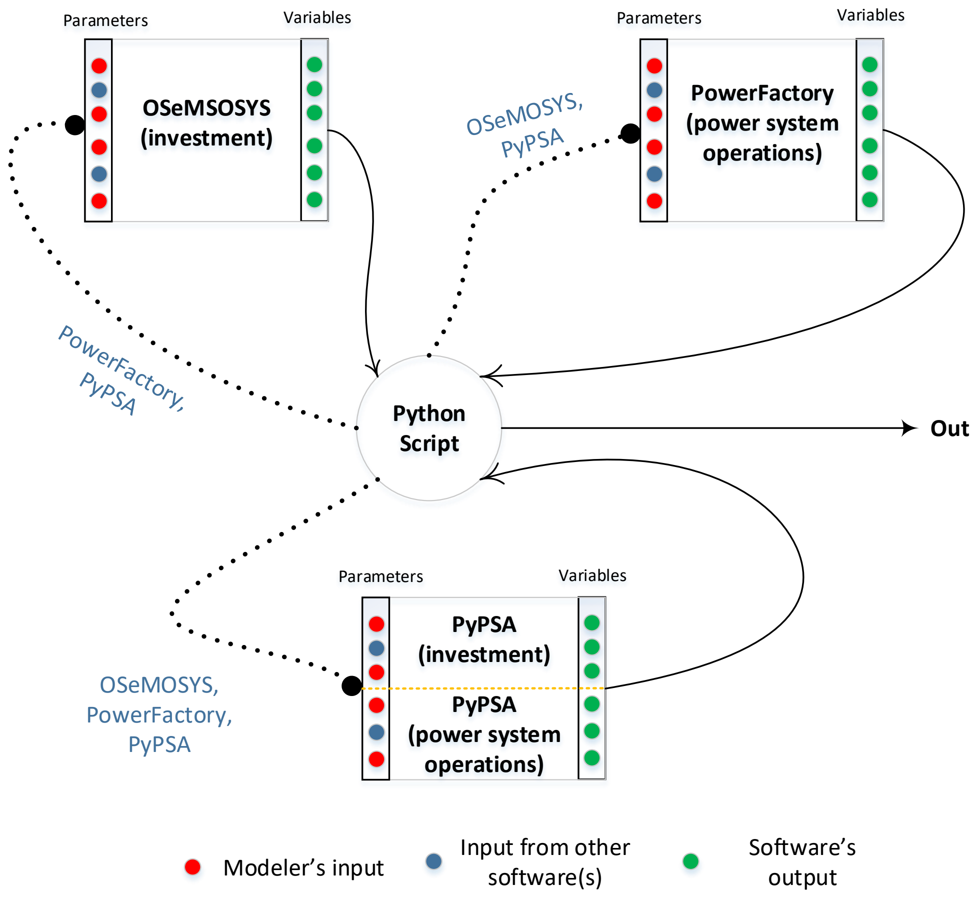


Figure 3.2: Representation of the bidirectional softlinking between the computational tools used in this thesis (OSeMOSYS, PowerFactory and PyPSA). The main Python script produces the physical exchange between parameters and variable results while checking for the convergence criterion against a pre-defined threshold.

neither created nor solved, unless the function is externally called. In this script, only the parameters (i.e. *params*) to be changed during the simulation run (blue dots in Fig. 3.2) must be specified as inputs to the function, while the rest of the parameters always remain as initially inserted in the **.csv** file (red dots in Fig. 3.2). Moreover, the variables (i.e. *vars*) to be extracted from the function (green dots in Fig. 3.2) are to be specified at the end of the function's body.

3.3.2. PowerFactory: Loaders, Executors and Extractors

PowerFactory is most commonly operated by means of a user interface. This interface allows to connect objects easily, assign properties to them, run simulations and visualize the results. While highly convenient, this way of operation does not meet the needs of the bidirectional softlinking method, as it requires a constant handling of the user. However, PowerFactory offers an advanced feature called **Scripting and Automation**, which allows the user to access PowerFactory via Python-developed commands. When run

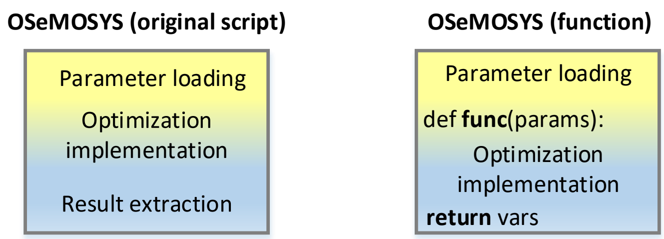


Figure 3.3: Representation of OSeMOSYS adaptation to fit the bidirectional softlinking scheme. The original OSeMOSYS script, a stand-alone file, is transformed into a module where the optimization problem is encapsulated by a function and can thus only be accessed through its input parameters and output variable results.

in this way, PowerFactory is considered to be operating in *engine mode*, in which the user has full external access to the software's functionalities.

To perform the coupling, some of these Python commands are grouped here in a Python module with its respective sets of functions. The three sets of functions are: 1) **Loaders**, which load parameters (both inserted by the user and subject to change during the simulation), 2) **Executors**, which run the functions in PowerFactory to perform the calculations needed and 3) **Extractors**, which retrieve the results after the calculations have been completed and load them to the Python workspace.

Fig. 3.4 shows an overview of the sequential implementation of these blocks. As observed in the figure, the loader loads to PowerFactory values changing during the simulation such as the installed capacity (variable result from OSeMOSYS), the fixed costs (directly proportional to the installed capacity), but also fixed parameters such as the minimum down-time constraints for a certain technology. After parameters have been loaded, the executor is called to run a unit commitment or an optimal power flow. Finally, the extractors are called to obtain the power plant schedule (when they will be on or off), their actual power output (which considers all commitment constraints) and their variable costs of operation (which depend on the actual power output). The reader is encouraged to look at Appendix A.2 for the code of specific functions contained in these sets such as **char_load**(*params*) from **Loaders** and **results**(*params*) from **Extractors**.

3.3.3. PyPSA: The Network container and its components

PyPSA's central component is the **Network**. This is an overall container in which all other components exist and where calculations are run, stored and can be retrieved from. After the **Network** component is called, it is possible to create other objects such as generators, storage units, loads, buses, stores and links. As observed in Fig. 3.5 and briefly mentioned in Subsection 2.4.2 , these objects have several parameters and variable results that can be monitored. For instance, generators can be either extendable (i.e. considered for the investment problem) or committable (i.e. considered for UCED operations). If generators are extendable, parameters such as their capital costs and variable costs must be specified. If they are committable, variable costs, start-up costs and inter-temporal constraints are taken into account instead. For

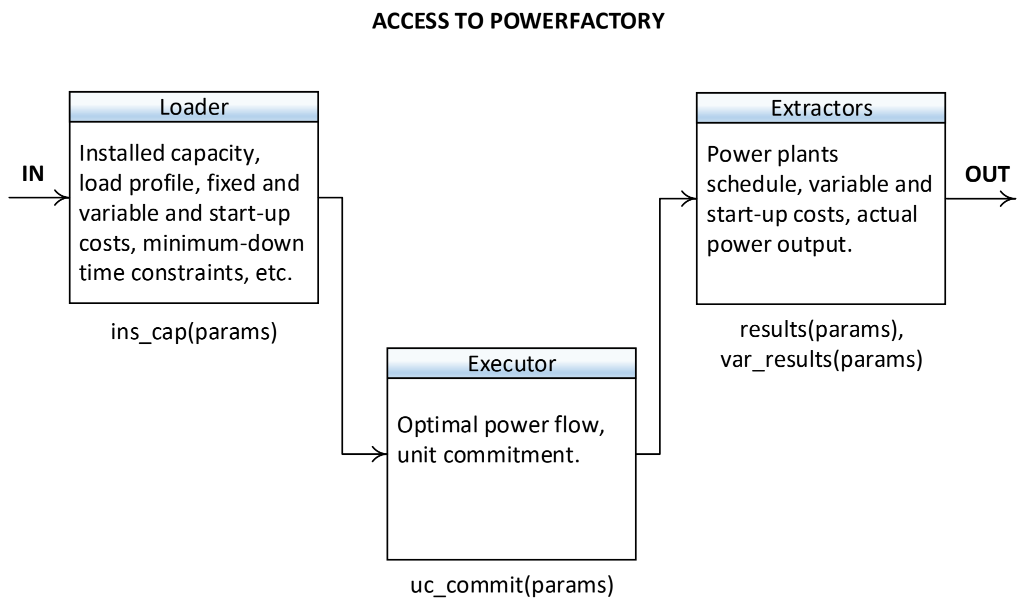


Figure 3.4: Scheme to access PowerFactory in engine mode (i.e. via Python commands while disregarding the user interface). For the bidirectional softlinking, the sets of functions group some of these commands in order to load parameters (e.g. installed capacity from OSeMOSYS), execute PowerFactory functions (e.g. unit commitment) and extract the calculation results (e.g. actual power output).

other components such as storage units, stores and links, the transition between the investment mode and the operation model is simply the capital cost.

Since PyPSA is Python-based, its integration in the bidirectional softlinking is simple. First, the Network component must be created. Then, the other components (generators, storage units, loads, grid connections, etc.) can be specified with their corresponding input parameters, e.g. those shown in Fig. 3.5. If these are several, a well-defined Python *for loop* can easily create the whole system within seconds without having to turn to sophisticated engine mode interfaces (i.e. PowerFactory). After the network is created, it is possible to solve the optimization problem by calling a linear optimal power flow function that is inherent to the Network component. When the problem is feasible and solved, multiple results can be accessed, again via the Network component. These results include the specific power output of all generators at every timestep, their start-up costs if they are comittable and their optimal power size if they are extendable. It is also possible to retrieve the state of charge of storage units, their output power at every timestep, and their optimal power sizes if they are extendable too. Several other variables that can be obtained from other components are specified in Fig. 3.5.

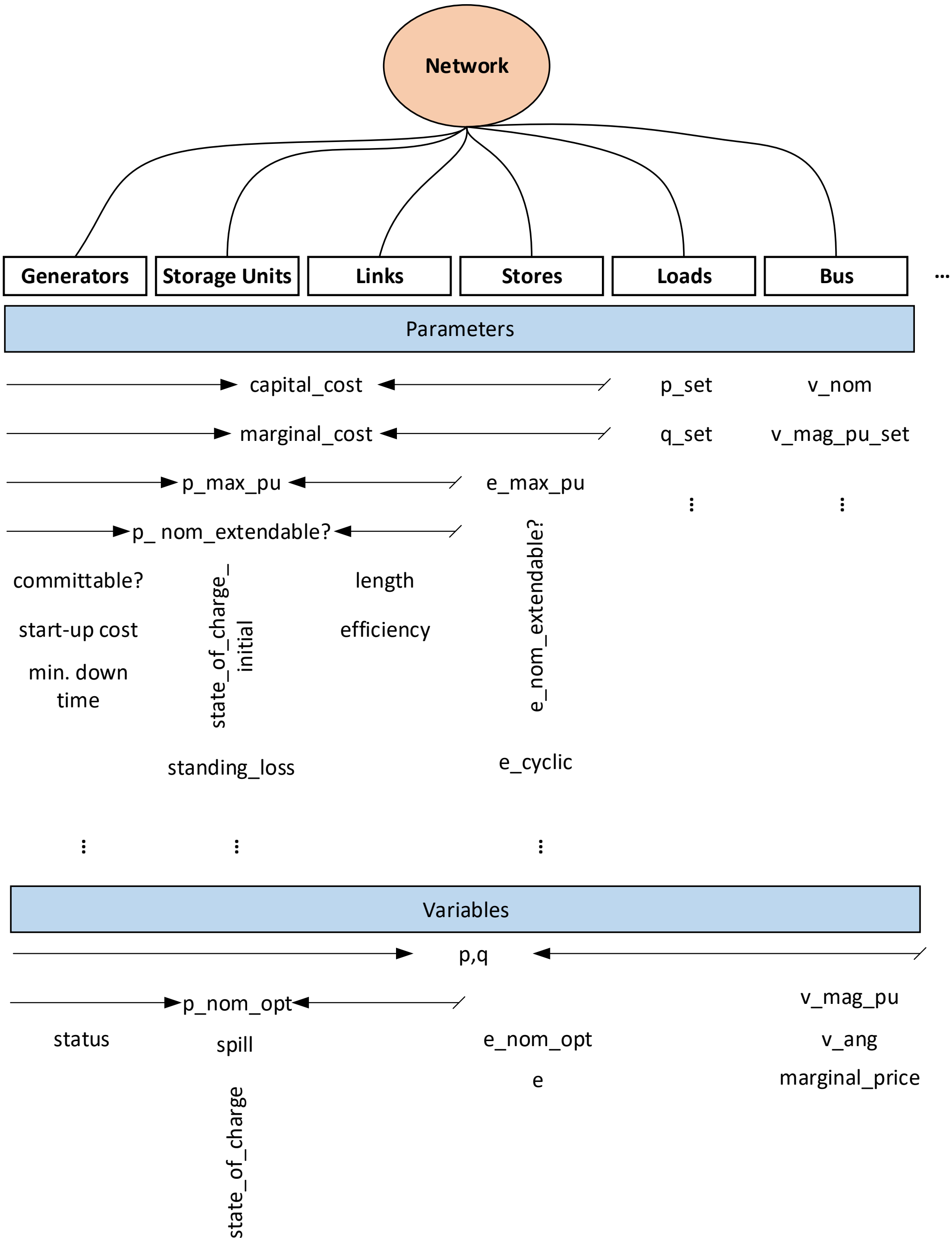


Figure 3.5: PyPSA's Network container and its components. Each component has several parameters and variable results available, which can be easily accessed via commands that exploit the object-oriented nature of Python. This makes the integration of PyPSA in the bidirectional softlinking a relatively simple task.

4
Simulation: Methods and results

4.1. Introduction

As discussed in Chapter 2 and remarked in Chapter 3, energy system optimization models (ESOMs) have two main limitations due to their typical use: 1) low-time resolution and lack of short-term power system operations equations. In Subsection 2.2.4 in particular, the concept of a *timeslice* was introduced. As mentioned there, the idea of *what* and *how many* timeslices are needed to improve the low-time resolution in relation to renewable energy integration is still an open research question. In this context, this chapter elaborates on a very specific technique based on a data clustering algorithm that helps choose timeslices based on the existing relation between power data points.

Furthermore, this chapter proposes three specific case studies to evaluate how short-term power system operations can be included in long-term investment models and how significant their influence is. Case study I consists of a two-generator power system in which the greatest part of the bidirectional softlinking is introduced in detail, developed and corroborated. Case study II then tests the same methodology used in the previous case study in a four-generator system in which the conditions are changed: renewable energy integration is increased, and the capital costs of technologies and their corresponding interest rates are varied. Finally, Case study III considers the entire Dutch fleet of conventional generators, it adds batteries to the system and allows renewable energy sources to become part of the investment solution. The latter case study is proposed as a way to bring the bidirectional softlinking closer to a real-size power system application.

4.2. Coarse time-resolution improvements: Representative timeslices

As highlighted in Subsection 2.2.4, one of ESOM's limitations is the low-time resolution that they use in order to lower the computational complexity of their simulations. As it was mentioned, this low-time resolution is achieved by selecting specific *timeslices* (single power data points with extended time duration) to represent load profiles and the availability of renewable energy sources. Among the vast amount of options to perform a methodological selection of these timeslices, the data clustering technique in Ref. [75] is the selected one, as it provides reliable results and it solves much faster than, for instance, the optimization approach in Ref. [73].

4.2.1. Method for selecting the representative timeslices

The data clustering technique in Ref. [75] relies on the Ward agglomerative hierarchical clustering algorithm, which allows to explore data at different levels of similarity, providing in this way a deeper insight into the relationships among samples [110]. This algorithm is described by the following four characteristics that must be specified for every technique in the hierarchical clustering family:

- **Initial condition**: Every point starts as a cluster itself; thus, there are initially n_{ct}^0 clusters.
- **Space and distance**: Clusters lie within an Euclidean space so that the distance between them is the L^2 norm for real vectors; that is, the Euclidean distance.
- **Linkage criterion**: For clusters to be agglomerated, they must share a certain similarity measure. Therefore, clusters are merged based on what cluster unions lead to the minimum point distance variance within the newly formed clusters. This similarity measure (or linkage criterion) is known as Ward due to its original author in Ref. [111].
- **Stoppage criterion**: A fixed number n_{ct} of clusters is pre-defined. That is, the algorithm stops when all the initial clusters (equivalent to all initial points) have been agglomerated into n_{ct} clusters.

What is a point/sample in the algorithm? For this, consider a timeseries that contains a normalized (i.e. by its maximum value) demand profile in a row vector $\overrightarrow{v_{1,8760}}^{dem}$, with full resolution. Then, assume this vector is reshaped as a matrix $V_{n_{dy},n_{hr}}^{dem}$, with $n_{dy} = 365$ (for the days in a year) and $n_{dy} \cdot n_{hr} \leq 8760$, where n_{hr} represents the amount of hours in a day, but that it is not necessarily mapped to $n_{hr} = 24$, for reasons explained hereinafter (thus, the inequality). Assume as well that other $\lambda - 1$ timeseries of equal size (e.g. normalized solar and wind profiles) are contained in similar row vectors and are equally reshaped as matrices. Finally, assume that the resulting matrices are concatenated into a single matrix: $V_{n_{dy},\lambda n_{hr}} = [V_{n_{dy},n_{hr}}^{dem}, V_{n_{dy},n_{hr}}^{solar}, V_{n_{dy},n_{hr}}^{wind}, ...]$. When $V_{n_{dy},\lambda n_{hr}}$ is passed to the algorithm, it is in fact interpreted as a column vector with n_{dy} data points or days (hence, $n_{ct}^0 = n_{dy}$), each with λn_{hr} coordinates.

Once the agglomerative hierarchical clustering algorithm has been applied to the data, it is necessary to interpret the information resulting from the n_{ct} clusters and how this translates to finding suitable *representative days* for the n_{dy} initial days. Ref. [75] proposes to choose a single point from each cluster so that this one represents all other samples within its corresponding cluster. Given that there are n_{ct} clusters and one representative sample per cluster, there are also n_{ct} representative days. With this information, it is possible to obtain the matrix $V_{n_{ct},\lambda n_{hr}}$, which represents the original one $V_{n_{dy},\lambda n_{hr}}$. Ultimately, the representative matrix can be dissaggregated to find the individual representative matrices for each timeseries; thus, the representative matrix for the demand profile, for instance, is $V_{n_{ct},n_{hr}}^{dem}$, which when unfolded, gives the representative timeseries $\overrightarrow{v_{1,n_{ts}}}^{dem}$, where its size corresponds to the number of timeslices $n_{ts} = n_{ct} \cdot n_{hr}$.

The next question is *how* to select one representative sample per cluster. Ref. [75] evaluates two potential candidates for this purpose: 1) the cluster centroid (i.e. the center of mass of the cluster) and 2) an actual sample that is the closest to the cluster centroid. After multiple experiments, Ref. [75] found that the second option provided a more accurate cluster representation for their dataset. Here, it is assumed that this conclusion also holds; thus, the closest point to each cluster centroid is chosen as a representative day. Another aspect to be decided is the duration of each selected representative day. Ref. [75] makes the duration of each representative day proportional to the cluster size in reference to the initial n_{dy} days. This means that if cluster i contains $n_{i\subset ct,s}$ samples, the duration of its representative day i is $d_i = \omega_i \cdot 8760$, where $\omega_i = (n_{i\subset ct,s}/n_{dy})$ is the weight or its constant of proportionality.

The description above regarding the agglomerative hierarchical clustering as well as the method proposed in Ref. [75] are visually summarized in Fig. 4.1 to help the reader have a better understanding. Furthermore, it is worth mentioning that the author found this algorithm extremely fast, powerful and accurate and thus kept exploring about possible existing software packages that could contain several other algorithms of this sort. This resulted in a successful search that it is herein shared if needed for furher research purposes: tsam - Time Series Aggregation Module (a Python package, available at https://github.com/FZJ-IEK3-VSA/tsam), supported by a scientific publication Ref. [112].

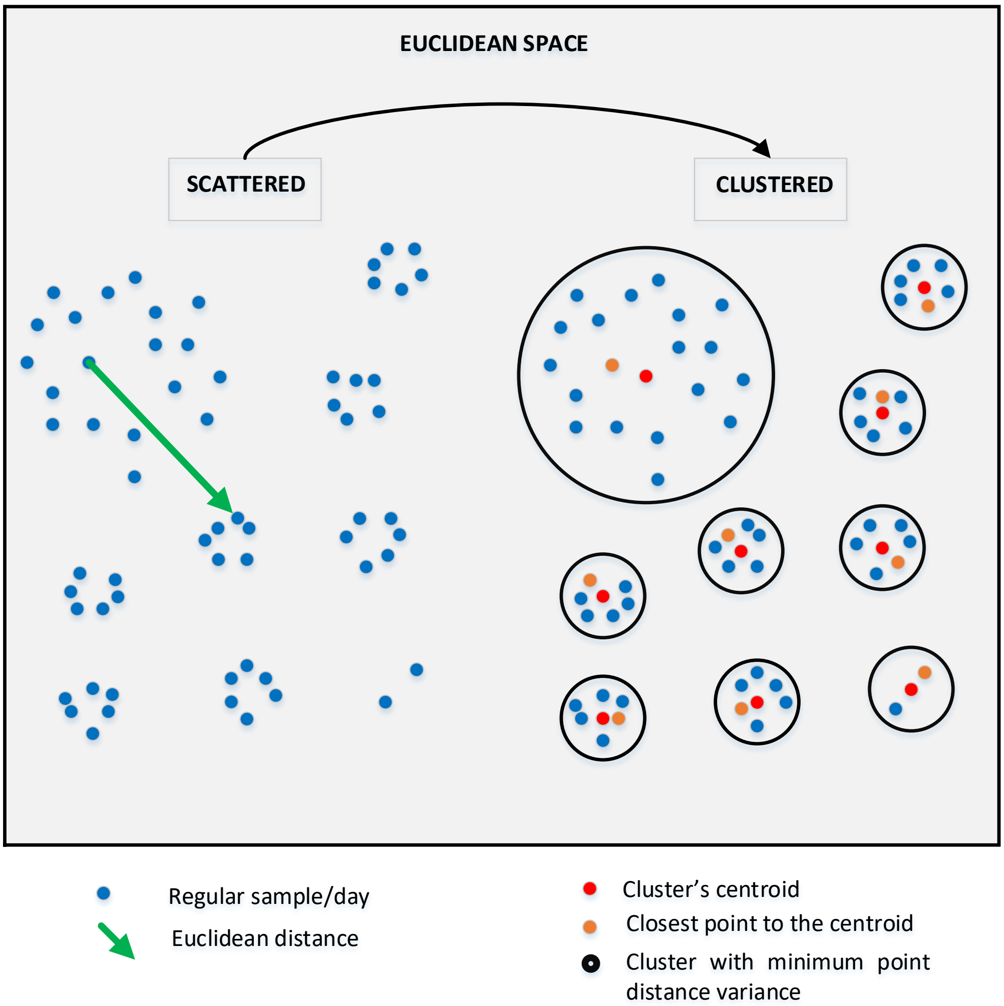


Figure 4.1: Representation of the Ward agglomerative hierarchical clustering with the considerations specified in Ref. [75]. These considerations include: measuring the distance between points based on the Euclidean distance, agglomerating clusters based on the unions that lead to the minimum distance variance within each new cluster; and finally, choosing the point closest to the cluster centroid as the representative data point for the whole cluster.

4.2.2. Heuristics to select the number of representative timeslices

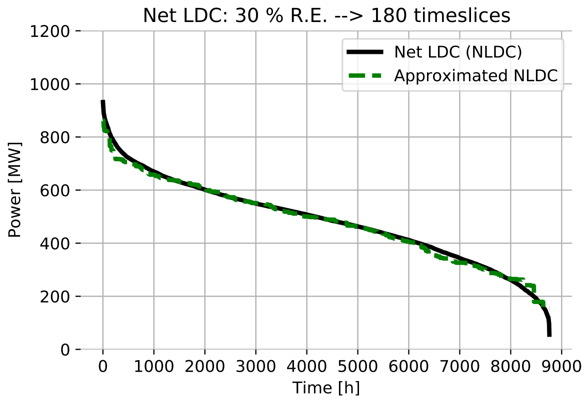
As explained in Subsection 4.2.1, the number of timeslices for a given timeseries is $n_{ts} = n_{ct} \cdot n_{hr}$, where $n_{dy} \cdot n_{hr} \leq 8760$, where $n_{dy} = 365$. To further reduce the computational complexity, n_{hr} is chosen to be $n_{hr} = 8$ in Ref. [75] so that each day is represented by 8 hours. These 8 hours are simply the original 24 hours split in blocks of 3 consecutive hours, as it is argued that this resolution is good enough to include renewable power fluctuations. The total number of representative days n_{ct} is then chosen according to several error metrics and a specific assessment based on the applications shown.

Here, the procedure applied is different. Instead, a maximum number of timeslices n_{ts}^{max} is first specified and the algorithm is then applied to both n_{ct} and n_{hr}. Several combinations of n_{ct} and n_{hr} are tested while a self-defined metric hereinafter specified is in charge of indicating how may timeslices n_{ts} there should be in total. The actual process is subsequently explained using a generic demand profile and its reconstructed counterpart as an example.

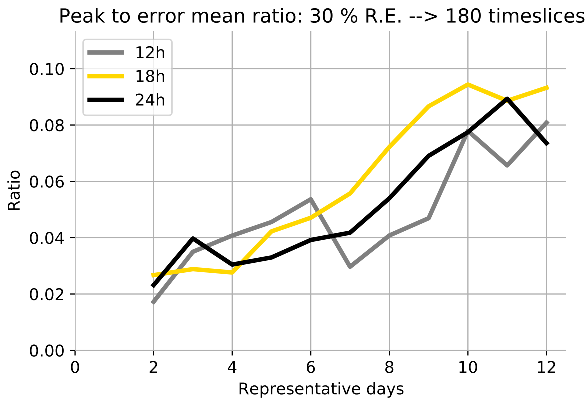
The first part of the process consists in obtaining a reconstructed demand profile $v_{1,8760}^{dem'}$, which can be implemented by repeating each representative timeslice u in $v_{1,n_{ts}}^{dem}$ by its duration $d_u = d_i \cdot d_j$ (see that $d_j = \omega_j \cdot 24$ is the duration of a representative hour j in a representative day i). Afterwards, it is necessary to sort this reconstructed profile to convert it to a reconstructed load duration curve $v_{1,8760}^{ldc'}$. The second step is to sort the original demand profile $v_{1,8760}^{dem}$ to obtain its respective load duration curve $v_{1,8760}^{ldc}$. The metric to assess the accuracy of the reconstructed load duration curve is the $peak\_to\_error\_mean$, which is in fact composed of two sub-metrics: 1) the mean of the absolute error $\overline{error_{abs}}$ and 2) the peak ratio $peak\_ratio$. The latter depicts how accurately the representative timeseries resembles the peak of the original timeseries, which is particularly important as the curve peak defines the amount of installed capacity in an energy mix. The metric can thus be fully expressed as:

$$peak\_to\_error\_mean = \frac{peak\_ratio}{\overline{error_{abs}}} = \frac{\max\left(v_{1,8760}^{ldc'}\right) / \max\left(v_{1,8760}^{ldc}\right)}{\overline{\left| v_{1,8760}^{ldc'} - v_{1,8760}^{ldc} \right|}} \tag{4.1}$$

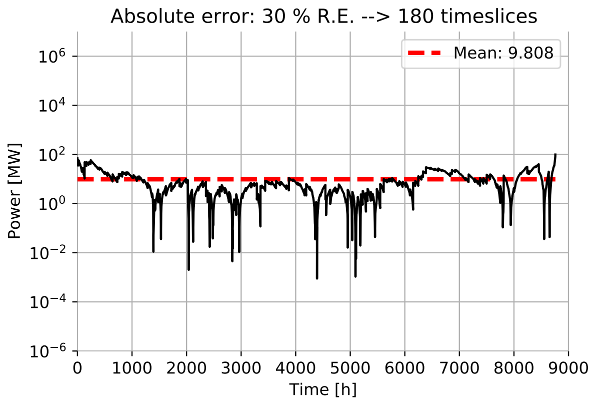
As observed, $peak\_to\_error\_mean$ achieves its maximum value when the trade-off between the peak representation and the mean of the absolute error is the highest (i.e. it is best to have a good peak representation while the mean of the errors between curves is also small). Next, it is assumed that the number of hours n_{hr} can be either 12, 18 or 24 hours. The maximum number of timeslices is set to $n_{ts}^{max} = 288$ (similar to Ref. [71]). For each n_{hr}, n_{ct} is varied from $n_{ct} = 2$ to $n_{ct} = n_{ts}^{max} / n_{hr}$. For the 1000 MW Dutch load duration curve with 30% renewable energy integration shown in Fig. 4.2a, the $peak\_to\_error\_mean$ is the highest at $n_{hr} = 18$ and $n_{ct} = 10$, as observed in Fig. 4.2b. This results in $n_{ts} = 180$ timeslices. Fig. 4.2a shows as well how the reconstructed demand profile $v_{1,8760}^{dem'}$ approximates the original demand profile $v_{1,8760}^{dem}$, while Fig. 4.2c shows how the error changes along the curves with a mean absolute error of $\overline{error_{abs}} = 9.808$ MW. Despite the error introduced, the reduction is remarkable: from 8760 timeslices to 180 timeslices, which is equivalent to a 97.945% reduction.



(a) Full resolution of the net load duration curve (NLDC) compared to the reconstructed NLDC. The latter NLDC is obtained by repeating each timeslice based on the duration of its representative hours and representative days.



(b) Peak-to-error-mean ratio. The yellow curve shows that the higest ratio is reached when the number of representative hours is $n_{hr} = 18$, while the number of representative days is $n_{ct} = 10$. This is equal to $n_{ts} = 180$ timeslices.



(c) Absolute error between the full resolution NLDC and the reconstructed NLDC at every timestep.

Figure 4.2: Demonstration of the Ward agglomerative hierarchical clustering algorithm and the self-defined heuristics to find the appropriate number of representative hours and representative days.

4.2.3. Representing the actual profile peak

Although the $peak\_to\_error\_mean$ metric defined in Subsection 4.2.2 explicitly prioritizes the representation of the curve peak $\max\left(v_{1,8760}^{ldc}\right)$, the algorithm may not select this specific power data point. However, the lack of this information in the ESOM would result in an underestimated installed capacity that once passed to the UCED model would most likely lead to an unfeasible problem. Thus, the peak must be enforced in the ESOM without disturbing the actual representative timeslices.

A way to include the peak is by defining a timeslice that has an extremely small weight and consequently a small duration too: $d_u = \omega_u \cdot 8760$, with e.g. $\omega_u = 1 \cdot 10^{-8}$. Afterwards, the height of the timeslice can be specified to be the curve peak $\max\left(v_{1,8760}^{ldc}\right)$. This way the timeslice becomes a spike, which represents an infinitesimal fraction of the energy under the curve.

4.3. Case study I: Two-generator system

4.3.1. Purpose and methodology

This case study is the first experiment that intends to answer the second specific research question in Section 1.6. The simplicity of this case study allows to pay attention to the development of the model coupling method (bidirectional softlinking) and the benefits and consequences that it carries. Specifically, the purpose of this case study is to understand the circumstances in which power plant short-term operations, i.e. their inter-temporal dynamic constraints in the UCED problem, could influence long-term investments. To this end, OSeMOSYS first decides the installed capacity of two power plants, each corresponding to a different technology. The circumstances assigned to the experiment are different degrees of flexibility in the power plants' UCED constraints and a relatively high level of renewable energy integration. Later in the process, PowerFactory evaluates three specific variables: actual variable costs of operation, the start-up costs and the actual energy output of the mix. Furthermore, a proposed bidirectional softlinking (BSL) is applied, which communicates to OSeMOYSYS all these three variable results in order to push the iterative process towards a lower total system cost. Finally, the result from the BSL is corroborated with a brute-force technique, which is computationally expensive but allows to keep linearity.

4.3.2. Description

Consider that a power system with two power plants must be built. The technologies to be included are nuclear and coal with financial parameters as per Table 2.2. Moreover, consider that these two technologies have two sorts of UCED parameters [113]: flexible and inflexible, as specified in Table 4.1. Flexible parameters are those that allow the plants to be more dynamic in such a way that their output or status (i.e. on/off) change more often. On the other hand, inflexible parameters are those that impose more stringent constraints on the plants' dynamics and thus force them to remain in a similar state for longer periods.

Table 4.1: Flexible and inflexible parameters related to the constraints minimum-down time, minimum-up time and minimum load. These values have been retrieved from Ref. [113], where a similar study of flexible vs. inflexible mixes was performed. Start-up costs (cold-start and warm-start) are the same for both cases.

Technology	Cold-start cost (EUR/MW)	Warm-start cost (EUR/MW)	Minimum-down time (h)		Minimum-up time (h)		Minimum load (p.u.)	
			Flex	Inflex	Flex	Inflex	Flex	Inflex
Nuclear	140	100	24	24	0.25	24	0.40	0.50
Coal	74	42	3	10	0.25	10	0.25	0.40
CCGT	45	33	0.50	6	0.25	6	0.30	0.50
COCGT	28.5	21	0.25	1	0.25	1	0.20	0.50

This two-generator power system must supply the load duration curve shown in Fig. 4.2. By applying the standard screen curve method explained in Subsection 2.2.3 or by inserting the power plants investment parameters in OSeMOSYS (with the needed modifications as expressed in Appendix A.1), it is possible to obtain the set of solutions that disregards the UCED constraints. The installed capacity for each technology is: i^{cap}_{nuc} = 243.94 MW and i^{cap}_{coal} = 687.55 MW. Consequently, their total fixed costs can be calculated: $c_g^{total\_fix} = \left(C^{acc}_{nuc} + C^{fix}_{nuc}\right) i^{cap}_{nuc} + \left(C^{acc}_{coal} + C^{fix}_{coal}\right) i^{cap}_{coal}$= 95.61 MEUR + 109.32 MEUR= 204.93 MEUR.

4.3.3. Unidirectional softlinking (USL): Capturing missing costs

As derived in Subsection 4.3.2, the total fixed costs are c^{total_fix} = 204.93 MEUR for the energy mix. Regarding the total variable cost of operation c^{total_var}, it is calculated by OSeMOSYS that $c^{total\_var,ose} = c^{var,ose}_{nuc} + c^{var,ose}_{coal}$ = 38.07 MEUR + 97.96 MEUR= 136.03 MEUR. However, it remains to compare these results with those obtained from the UCED model in order to evaluate whether its constraints have a significant influence over the costs. To this end, the methodology specified in the unidirectional softlinking (USL) block in Fig. 4.3 is followed for the flexible and inflexible parameters. Note that the total fixed costs (capital costs plus fixed costs) are not discussed for USL, as installed capacities are not subject to change and so these costs are not either.

In this context, Fig. 4.4 offers a comparison between the variable cost of operation $c^{total\_var,ose}$ = 136.03 MEUR calculated by OSeMOSYS and the actual variable cost of operation $c^{total\_var,pf}$ obtained from PowerFactory. For the flexible parameters, Fig. 4.4a shows that $c^{total\_var,pf}$ = 151.07 MEUR, which represents a mismatch of 15.04 MEUR with respect to OSeMOSYS. Additionally, there are $c^{total\_su,pf}$ = 4.04 MEUR in total start-up costs. This raises the total mismatch to 19.08 MEUR. Fig. 4.4b then shows the power plants' perspective. As observed, the use of the nuclear power plant is overestimated by OSeMOSYS, while the use of the coal power plant is underestimated. The most significant misrepresentation is for the coal power plant, with 26.93 MEUR of underestimated costs.

The process is subsequently repeated for the inflexible parameters. Fig. 4.4c shows that the start-up costs $c^{total\_su,pf}$ = 4.51 MEUR are only slightly higher in comparison to the flexible parameters. Nonetheless, the variable cost of operation increases even more, $c^{total\_var,pf}$ = 166.67 MEUR. Taking these two costs into consideration, the total mismatch between OSeMOSYS and PowerFactory is even greater: 35.15 MEUR. With regards to the power plants, Fig. 4.4d shows a similar trend as before: nuclear is overestimated, while coal is underestimated. However, the individual mismatches between generators are much more noticeable than with the flexible parameters. The case of the coal power plant is particularly important as the mismatch between OSeMOSYS and PowerFactory is 52.18 MEUR for this single generator.

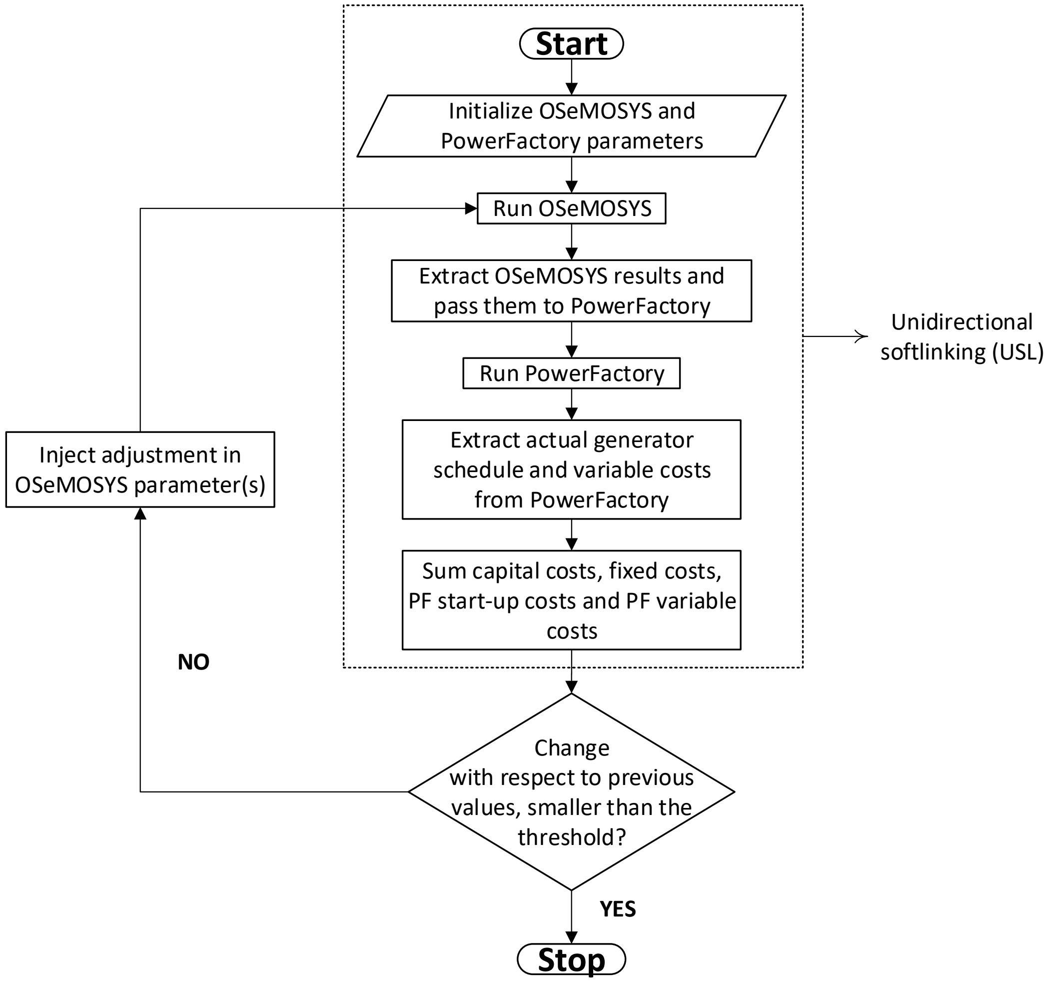
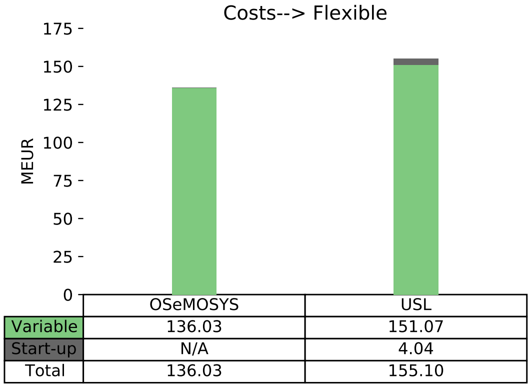
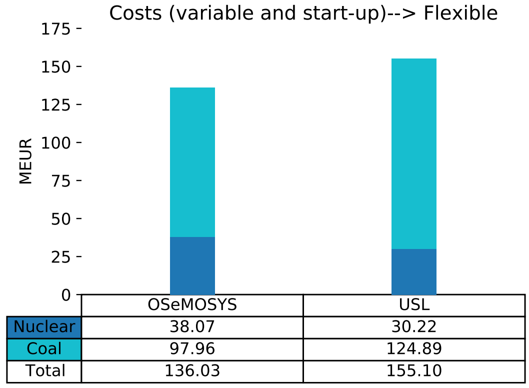


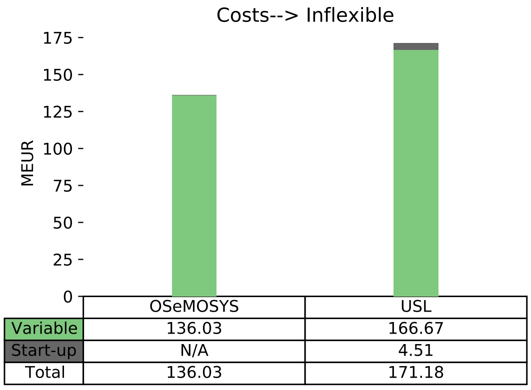
Figure 4.3: Flowchart of the bidirectional softlinking (BSL) coupling method between OSeMOSYS and PowerFactory. The process consists in an unidirectional softlinking (USL; a subset of BSL as illustrated by the dashed block), whose results are fed back to OSeMOSYS in each iteration via designed adjustments. The value to be tracked in this specific BSL is the total system cost, composed of capital costs, fixed costs, actual variable costs (calculated by PowerFactory) and start-up costs.



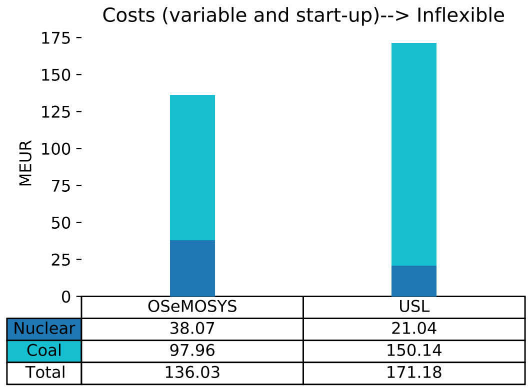
(a) Total variable cost of operation and total start-up cost, with flexible parameters. The total mismatch between OSeMOSYS and PowerFactory amounts to 19.08 MEUR.



(b) Sum of the total variable cost of operation and the total start-up cost for each power plant with flexible parameters. The use of nuclear is overestimated by OSeMOSYS, while the use of coal underestimated.



(c) Total variable cost of operation and total start-up cost, with inflexible parameters. The total mismatch between OSeMOSYS and PowerFactory amounts to 35.15 MEUR.



(d) Sum of the total variable cost of operation and the total start-up cost for each power plant with inflexible parameters. The use of nuclear is overestimated by OSeMOSYS, while the use of coal underestimated.

Figure 4.4: Comparison between the total variable cost of operation calculated by OSeMOSYS and the one calculated by PowerFactory while considering the flexible and inflexible parameters shown in Table 4.1. The total start-up cost, which is only possible to obtain from PowerFactory, is additionally shown. The misrepresented costs are greater for the inflexible parameters (i.e. 35.15 MEUR) in comparison to the flexible parameters (i.e. 19.08 MEUR). Furthermore, the coal power plant is the least accurately represented in both cases with 52.18 MEUR and 26.93 MEUR in mismatches, respectively.

4.3.4. Bidirectional softlinking (BSL): Creating feedback loops

As indicated in Subsection 4.3.3, including the UCED equations significantly increases the total variable cost of operation of the energy mix and introduces other costs, e.g. costs for starting the power plants. As important as this information is in order to have a much better representation of the costs for long-term investment decisions, it is equally relevant to wonder whether this information can be exploited to reach lower costs. To this end, the bidirectional softlinking method is applied (see Fig. 3.1 and Subsection 3.2.3 for further reference about this method) to influence the decisions taken by OSeMOSYS. As observed in

Fig. 4.3, this can be implemented via a feedback loop that returns to OSeMOSYS after obtaining the total costs from the unidirectional softlinking (USL) process. The feedback loop is an *adjustment* (see Subsection 3.2.3 where this term was first introduced) performed in one or several parameters in OSeMOSYS, which leads to a *possibly* different solution in every iteration. Iterations stop when solutions become too similar within a given range.

Here, two adjustments are proposed with the intention of achieving lower costs via communicating to OSeMOSYS the actual operation of the power plants obtained by PowerFactory:

- **Adjustment 1: Constraint on the energy output and injection of the start-up costs**

  Definition:

  In OSeMOSYS, the energy output of each technology is optimally tuned to supply demand. This output is originally unrestricted as power-plant inter-temporal equations are not included and thus OSeMOSYS cannot take into account essential constraints such as e.g. minimum-down time. This adjustment therefore consists in first obtaining the actual energy produced by each technology based on the UCED problem in PowerFactory. Then, this information is passed as an upper limit to the energy output of the corresponding technologies in the investment model. The other part of the adjustment implements a similar procedure with the start-up costs, which are also only obtainable from PowerFactory. Subsequently, these costs are transformed and injected back in OSeMOSYS in a different form, i.e. as fixed costs (see Ref. [56], where this has been similarly done to start-up costs). Based on this methodology, the adjustment imposes a new constraint and injects new start-up costs in every iteration. Hereinafter, this adjustment is called *Energy limit and SU costs*.

  Mathematical description:

  For the first part of the adjustment regarding the constraint on the energy output, consider the single-period optimization CGEP in Subsection 2.2.3 for generator g, with an additional constraint:

$$\sum_{b} p_{g,b} \cdot d_b \leq E_g^{max}, \tag{4.2}$$

  where E_g^{max} is an upper limit on the energy produced by generator g. This part of the adjustment simply consists in updating $E_g^{max,ose_{(k+1)}}$ in OSeMOSYS at iteration $k+1$ with the energy $\sum_{b\subset T^{pf}} p_{g,b}^{pf_{(k)}} \cdot d_b$ calculated by PowerFactory at iteration k. Mathematically, this can expressed as follows:

$$\sum_{b\subset T^{ose}} p_{g,b}^{ose_{(k+1)}} \cdot d_b \leq E_g^{max,ose_{(k+1)}} = \sum_{b\subset T^{pf}} p_{g,b}^{pf_{(k)}} \cdot d_b. \tag{4.3}$$

  For the second part of the adjustment, consider again the single-period optimization CGEP in Subsection 2.2.3 for generator g, at iteration $k+1$:

$$c_{sys}^{total,ose_{(k+1)}} = \left(C_g^{acc} + C_g^{fix_{(k+1)}}\right) i_g^{cap,ose_{(k+1)}} + \sum_b C_g^{var} p_{g,b}^{ose_{(k+1)}} \cdot d_b, \tag{4.4}$$

where $C_g^{fix_{(k+1)}} = C_g^{fix} + C_g^{sup} \cdot \sum_{b \subset T^{pf}} s_{g,b}^{up,pf_{(k)}}$. Here, $C_g^{sup} \cdot \sum_{b \subset T^{pf}} s_{g,b}^{up,pf_{(k)}}$ represents the resulting total start-up cost per installed capacity of g (that is, the term has been normalized by $I_g^{cap,ose_{(k)}}$), obtained from PowerFactory at iteration k.

- **Adjustment 2: Re-scale of the variable cost parameter C^{var}**

Definition:

The variable cost parameter C^{var} is essential to calculate the energy output of a given technology as well as its installed capacity. This parameter is a part of both the central generation expansion planning (CGEP) problem in Subsection 2.2.3 and the UCED problem in Subsection 2.3.1. This inherent connection between these two problems enables a plausible way to make them exchange information. The base of this adjustment thus consists in calculating the difference between the variable cost of operation per technology obtained by OSeMOSYS and the actual variable cost of operation provided by PowerFactory. This difference at every time block is re-expressed as a constant of proportionality by mathematical means described below. This constant is then used to re-scale the variable cost parameter C^{var} in OSeMOSYS in the next iteration. At every iteration, a new constant of proportionality is produced, which solely depends on the aforementioned differences between the two softwares and it is hence no subject to any specific range. Hereinafter, this adjustment is called *Variable cost re-scale.*

Mathematical description:

The aim of this adjustment is to re-scale the variable cost parameter $C_g^{var_{k+1}}$ to be inserted in OSeMOSYS at iteration $k+1$ in terms of the original C_g^{var}. The re-scale consists of a simple constant of proportionality γ and the following mathematical relation:

$$C_g^{var_{(k+1)}} = \gamma \cdot C_g^{var}. \tag{4.5}$$

The selection of γ, however, requires a more elaborated process, which is explained next. When observing the single-period optimization CGEP in Subsection 2.2.3 and the UCED model in Subsection 2.3.1, it is possible to note that the term $C_g^{var} \sum_b p_{g,b} \cdot d_b$ for each specific generator g is shared by the objective function of both problems. When in iteration k, OSeMOSYS calculates its solution $p_{g,b}^{ose_{(k)}}$ while PowerFactory proceeds to do the same, $p_{g,b}^{pf_{(k)}}$, after it receives the energy mix. To represent the relation between these two parameters, another constant of proportionality $q_b^{(k)}$ is introduced, allowing therefore to establish the following equation (note that the right hand side of Eq. (4.6) also sums over $b \subset T^{ose}$ as the time resolution of the PowerFactory's output is adapted to fit the resolution of OSeMOSYS):

$$\sum_{b \subset T^{ose}} p_{g,b}^{ose_{(k)}} \cdot d_b = \sum_{b \subset T^{ose}} q_{g,b}^{(k)} p_{g,b}^{pf_{(k)}} \cdot d_b \tag{4.6}$$

Nonetheless, this system is under-determined and cannot be solved for $q_{g,b}$. Hence, another approach is to assume instead the simple relation between the power variables (i.e. $q_{g,b} = p_{g,b}^{ose_{(k)}} / p_{g,b}^{pf_{(k)}}$) and weight their contributions to Eq. (4.6) based on how much energy $p_{g,b}^{ose_{(k)}} \cdot d_b$ is produced in each block b with respect to the total energy $\sum_b p_{g,b}^{ose_{(k)}} \cdot d_b$. This can be formally expressed as:

$$q_{g,b}^{(k)} = \begin{cases} p_{g,b}^{ose_{(k)}} / p_{g,b}^{pf_{(k)}} & p_{g,b}^{pf_{(k)}} \neq 0 \\ 10 & p_{g,b}^{pf_{(k)}} = 0 \end{cases} \qquad \omega_{g,b}^{(k)} = \begin{cases} p_{g,b}^{ose_{(k)}} / \sum_b p_{g,b}^{ose_{(k)}} & \sum_b p_{g,b}^{ose_{(k)}} \neq 0 \\ 0 & p_{g,b}^{pf_{(k)}} = 0 \end{cases} \tag{4.7}$$

where $q_{g,b}$ is set arbitrarily large (e.g. 10) for a factor if $p_{g,b}^{pf_{(k)}} = 0$ and $\omega_{g,b}$ is not considered if $\sum_b p_{g,b}^{ose_{(k)}} = 0$. Finally, this derivation is related back to the $k+1$ iteration via γ, as the overall idea behind the feedback is that OSeMOSYS approximates PowerFactory's results, i.e. $p_{g,b}^{ose_{(k+1)}} = p_{g,b}^{pf_{(k)}}$. With this in mind, Eq. (4.5) becomes:

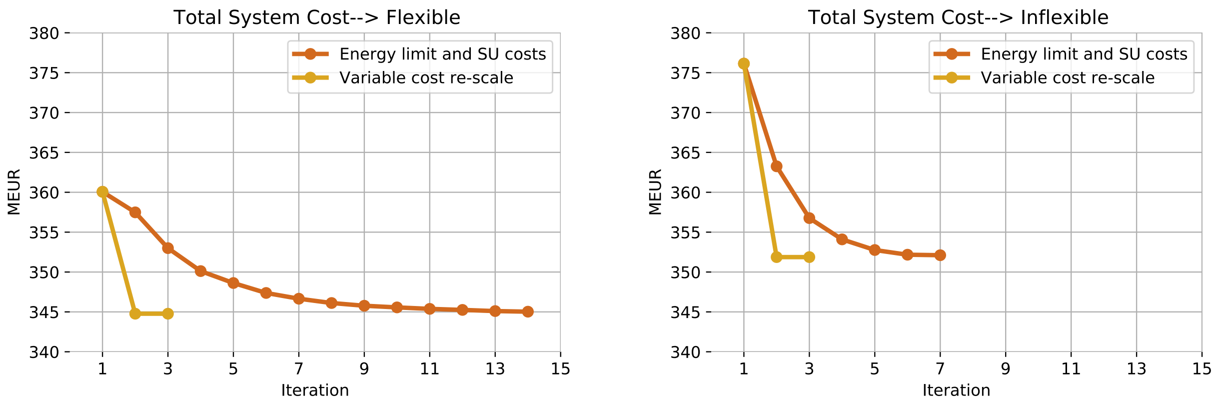
$$C_g^{var_{(k+1)}} = \sum_b \omega_{g,b}^{(k)} q_{g,b}^{(k)} \cdot C_g^{var}. \tag{4.8}$$

When revisiting the relation $q_{g,b}^{(k)} = p_{g,b}^{ose_{(k)}} / p_{g,b}^{pf_{(k)}}$, it is possible to observe that $q_{g,b}^{(k)} > 1$ when $p_{g,b}^{ose_{(k)}} > p_{g,b}^{pf_{(k)}}$. This increases $C_g^{var_{(k+1)}}$, leaving OSeMOSYS with most likely only two options, which it implements either separately or simultaneously: 1) decrease the installed capacity of g and 2) decrease the operation of g. On the other hand, when $p_{g,b}^{ose_{(k)}} < p_{g,b}^{pf_{(k)}}$, then $q_{g,b}^{(k)} < 1$, inducing OSeMOSYS to take contrary actions: 1) increase the installed capacity of g and 2) increase the operation of g. Therefore, the mismatch between the long-term model and the short-term model is captured by this relation, thus fulfilling the purpose established at first.

Comparison between adjustments:

The effectiveness of the adjustments introduced above in reducing costs is forthwith compared. First note that in contrast to the unidirectional softlinking (USL), the total fixed costs (capital costs and fixed costs) are considered here, as installed capacities do change in every iteration due to the adjustments injected in OSeMOSYS parameters. For further clarification, the initial solution of the total system cost is the one obtained by the USL at iteration $k=0$: $c_{sys}^{total,usl_{(0)}} = c^{total\_fix,ose_{(0)}} + c^{total\_var,pf_{(0)}} + c^{total\_su,pf_{(0)}}$. For the case with flexible parameters, $c_{sys,flex}^{total,usl_{(0)}}$= 204.93 MEUR + 151.07 MEUR + 4.04 MEUR= 360.04 MEUR. On the other hand, for the inflexible parameters, $c_{sys,inflex}^{total,usl_{(0)}}$= 204.93 MEUR + 166.67 MEUR + 4.51 MEUR= 376.11 MEUR.

Fig. 4.5 shows how both adjustments influence the total system cost during the course of the bidirectional softlinking (BSL) process. When the system has flexible parameters, it can be observed in Fig. 4.5a that Adjustment 1, *Energy limit and SU costs*, pushes the total system cost to $c^{total,bsl_{(14)}}_{sys,flex} \approx 345$ MEUR after $k = 14$ iterations. On the other hand, Adjustment 2, *Variable cost re-scale*, drives the total system cost to the same value $c^{total,bsl_{(3)}}_{sys,flex} \approx 345$ MEUR after only $k = 3$ iterations. When the system has inflexible parameters, on the other hand, it can be observed in Fig. 4.5b that *Energy limit and SU costs* moves the total system cost to $c^{total,bsl_{(7)}}_{sys,flex} \approx 352$ MEUR in $k = 7$ iterations, while *Variable cost re-scale* reaches a similar result, $c^{total,bsl_{(3)}}_{sys,flex} \approx$ 352 MEUR, in only $k = 3$ iterations again.



(a) *Energy limit and SU costs* vs. *Variable cost re-scale* for power plants with flexible parameters. Both adjustments lower the total system cost from $c^{total,usl_{(0)}}_{sys,flex} = 360.04$ MEUR to $c^{total,bsl}_{sys,flex} \approx 345$ MEUR. However, *Variable cost re-scale* does it in fewer iterations, i.e. $k = 3$, vs. the $k = 14$ iterations needed by *Energy limit and SU costs*.

(b) *Energy limit and SU costs* vs. *Variable cost re-scale* for power plants with inflexible parameters. Both adjustments lower the total system cost from $c^{total,usl_{(0)}}_{sys,inflex} = 376.11$ MEUR to $c^{total,bsl}_{sys,inflex} \approx 352$ MEUR. However, *Variable cost re-scale* does it in fewer iterations, i.e. $k = 3$, vs. the $k = 7$ iterations needed by *Energy limit and SU costs*.

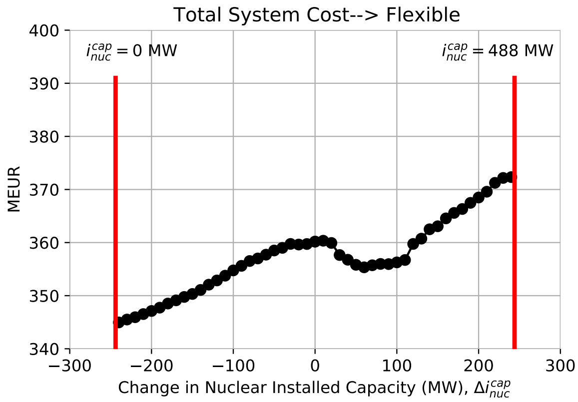
Figure 4.5: Comparison of the effectiveness between Adjustment 1, *Energy limit and SU costs*, and Adjustment 2, *Variable cost re-scale*, for cases in which the energy mix has flexible and inflexible parameters. For the two-generator system being considered, both options prove to be equally effective in lowering the total system costs; however, *Variable cost re-scale* proves to be the faster one.

4.3.5. Optimal solution: Space sweep

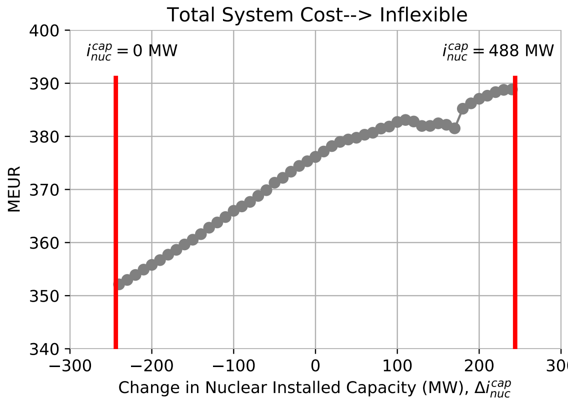
In Subsection 4.3.4, it was indicated that both adjustments manage to lower the total system cost. In a system with flexible parameters, both adjustments lead to $c^{total,bsl}_{sys,flex} \approx 345$ MEUR. On the other hand, both adjustments result in $c^{total,bsl}_{sys,inflex} \approx 352$ MEUR in a system with inflexible parameters. The aim of this part of the thesis is to validate whether these values obtained via the bidirectional softlinking are in fact optimal values. In order to do so, it would be necessary to solve the full ESOM-UCED optimization problem using a non-linear solver to find the true optimal values and then compare them. Alternatively, a co-optimization problem as shown in Subsection 3.2.1 can be solved, but installed capacities would need to be discrete and the resulting optimal values may be dependent on the sizes assigned to each power plant. Another option is to sweep the decision space, i.e. slowly increase and decrease installed capacities of each technology to observe which specific mix produces the lowest total system cost. Although this option is computationally expensive, the fact that only two generators are being considered still justifies its use.

As it was also shown in Subsection 4.3.2, the installed capacities in the energy mix are distributed as follows: $i_{nuc}^{cap,ose} = 243.94$ MW and $i_{coal}^{cap,ose} = 687.55$ MW. As i_{nuc}^{cap} is initially the smallest, the easiest option to sweep the space is thus to decrease this value to $i_{nuc}^{cap} = 0$ and then progressively increase it to a value where the trend of the total system cost becomes clear. Here, the final value adopted is $i_{nuc}^{cap} = 2 \cdot i_{nuc}^{cap,ose} \approx 488$ MW. More specifically, at every increment Δi_{nuc}^{cap}, the unidirectional softlinking (USL) process shown in Fig. 4.3 is repeated to obtain the total system cost for flexible and inflexible parameters: $c_{sys,flex}^{total}$ and $c_{sys,inflex}^{total}$, respectively. Note that for every step Δi_{nuc}^{cap}, the installed capacity of coal becomes $i_{coal}^{cap} = i_{coal}^{cap,ose} - \Delta i_{nuc}^{cap}$, as the total installed capacity must be preserved.

Fig. 4.6a shows the case for the flexible parameters. In this figure, it can be observed that the lowest value $c_{sys,flex}^{total} \approx 345$ MEUR is reached at $\Delta i_{nuc}^{cap} = -i_{nuc}^{cap,ose} = -243.94$ MW (equivalent to not having a nuclear power plant at all, i.e. $i_{nuc}^{cap} = 0$). Moreover, although there exists a valley around $\Delta i_{nuc}^{cap} = [0, 100]$ MW, $c_{sys,flex}^{total}$ is mostly monotonically increasing, revealing that a value lower than $c_{sys,flex}^{total} \approx 345$ MEUR is not expected beyond $i_{nuc}^{cap} \approx 488$ MW. A similar behavior is observed for a system with inflexible parameters shown in Fig. 4.6b. Here, the lowest total system cost $c_{sys,flex}^{total} \approx 352$ MEUR is also reached at $\Delta i_{nuc}^{cap} = -i_{nuc}^{cap,ose} = -243.94$ MW (i.e. no nuclear power plant, $i_{nuc}^{cap} = 0$), while $c_{sys,flex}^{total}$ mostly monotonically increases when nuclear is increased. With these observations in mind, it can thus be concluded that the true optimal value for flexible and inflexible parameters are the ones reached by the adjustments introduced in Subsection 4.3.4: $c_{sys,flex} \approx 345$ MEUR and $c_{sys,inflex} \approx 352$ MEUR.



(a) Space sweep for power plants with flexible parameters. The lowest total system cost is reached at $c_{sys,flex}^{total} \approx 345$ MEUR, when there is no nuclear power plant, i.e. $\Delta i_{nuc}^{cap} = -243.94$ MW or $i_{nuc}^{cap} = 0$ MW.



(b) Space sweep for power plants with inflexible parameters. The lowest total system cost is reached at $c_{sys,inflex}^{total} \approx 352$ MEUR, when there is no nuclear power plant, i.e. $\Delta i_{nuc}^{cap} = -243.94$ MW or $i_{nuc}^{cap} = 0$ MW.

Figure 4.6: Space sweep for the total system cost in a system with flexible and inflexible parameters. The space sweep consists in changing the initially calculated MW nuclear installed capacity, $i_{nuc}^{cap,ose} = 243.94$, in steps, Δi_{nuc}^{cap}. This happens while performing the reverse operation to coal, $i_{coal}^{cap} = i_{coal}^{cap,ose} - \Delta i_{nuc}^{cap}$, in order to keep a fixed total installed capacity. For every Δi_{nuc}^{cap} step, the unidirectional softliking (USL) method shown in Fig. 4.3 is applied.

4.4. Case study II: Four-generator system with different annualized capital costs

4.4.1. Purpose and methodology

This case study is the second experiment that intends to answer the second specific research question in Section 1.6. Similar to Case study I in Section 4.3, the purpose of this case study is to understand the circumstances in which power plant short-term operations, i.e. their inter-temporal dynamic constraints in the UCED problem, could influence long-term investments. However, this case study applies the methodology used above to a more diverse set of investment scenarios in which the penetration of renewable energy is increased, and the capital cost of technologies vary, as well as the interest rates at which such capital is obtained. This latter part of the experiment is particularly relevant, as the capital cost of a given technology can in fact significantly vary depending on the taxes, regulations and risk perception of a given region.

4.4.2. Description

The adjustments introduced in Subsection 4.3.4 offer promising results. As it was corroborated by space sweeping in Subsection 4.3.5, they push the total system cost to its lowest possible value for both flexible and inflexible parameters. However, no general conclusions can yet be drawn as it is necessary to test these adjustments in bigger systems, with a different range of costs, a variety of technologies, and higher levels of renewable energy integration. To implement this while keeping a reasonable computational time, two more generators are added: combined-cycle gas turbine (OCGT) and combustion open cycle gas turbine (COCGT). Furthermore, the system is tested at 30 % and 70 % of renewable energy integration.

The UCED parameters of the newly added power plants are given in Table 4.1 and all financial parameters, except the capital cost C^{cc}, are as per Table 2.2. In this case study, the capital cost C^{cc} is varied based on the argument that C^{cc} changes drastically depending on the taxes and regulations of the country in which a given power plant is to be built. For example, the capital cost of a nuclear power plant in Korea is estimated to be C^{cc}_{nuc} = 2021 USD/kW, while in Hungary this cost is as high as C^{cc}_{nuc} = 6215 USD/kW [10]. Similar differences can be found in the capital cost of other technologies, as documented in Ref. [10]. The methodology here adopted thus consists in creating cases in which different combinations of capital cost vectors $\vec{C^{cc}} = [C^{cc}_{nuc}, C^{cc}_{coal}, C^{cc}_{ccgt}, C^{cc}_{cocgt}]$ are assigned to the energy mix. More specifically, the capital cost of each technology is either the minimum, average or maximum of its range, as given by Ref. [10].

With the aforementioned methodology, the number of cases n_{cases} would grow to $n_{cases} = 3^4 = 81$, as there are 4 technologies with 3 options each. This is again too computationally expensive so the number of cases is reduced to only $n_{cases} = 6$. These are the cases in which nuclear and COCGT (baseload and high peaker, respectively) are either very cheap or very expensive, but this condition is mutually exclusive (e.g. they cannot be cheap simultaneously). Moreover, coal and CCGT (load-follower and peaker, respectively) are allowed to take either minimum and maximum values (also mutually exclusive) or have average values at the same time.

Finally, as these six capital cost vectors $\vec{C^{cc}}$ are ultimately annualized to obtain the annualized capital cost vectors $\vec{C^{acc}}$, another factor to consider is the discount rate R demanded by investors, which varies depending on the investors' perception of risk. For instance, $R = 0.07$ usually fits investments in deregulated markets, while $R = 0.10$ is more appropriate for risky environments [67]. After applying these two different discount rates, $R = 0.07$ and $R = 0.10$, the final number of cases grows to $n_{cases} = 12$. All cases are summarized in Table 4.2 and Table 4.3.

Table 4.2: Cases 1,2,3,4,5 and 6 for a varying annualized capital cost C^{acc} for the four technologies considered: Nuclear, Coal, CCGT and COCGT.

Cases	C^{acc} (kEUR/MW)			
	Nuclear	Coal	CCGT	COCGT
1	96.53	45.74	77..91	56.39
2	96.53	117.01	61.71	56.39
3	96.53	172.54	378.90	56.39
4	332.02	45.74	77.91	30.22
5	332.02	117.01	61.71	30.22
6	332.02	172.54	37.90	30.22

Table 4.3: Cases 7,8,9,10,11 and 12 for a varying annualized capital cost C^{acc} for the four technologies considered: Nuclear, Coal, CCGT and COCGT.

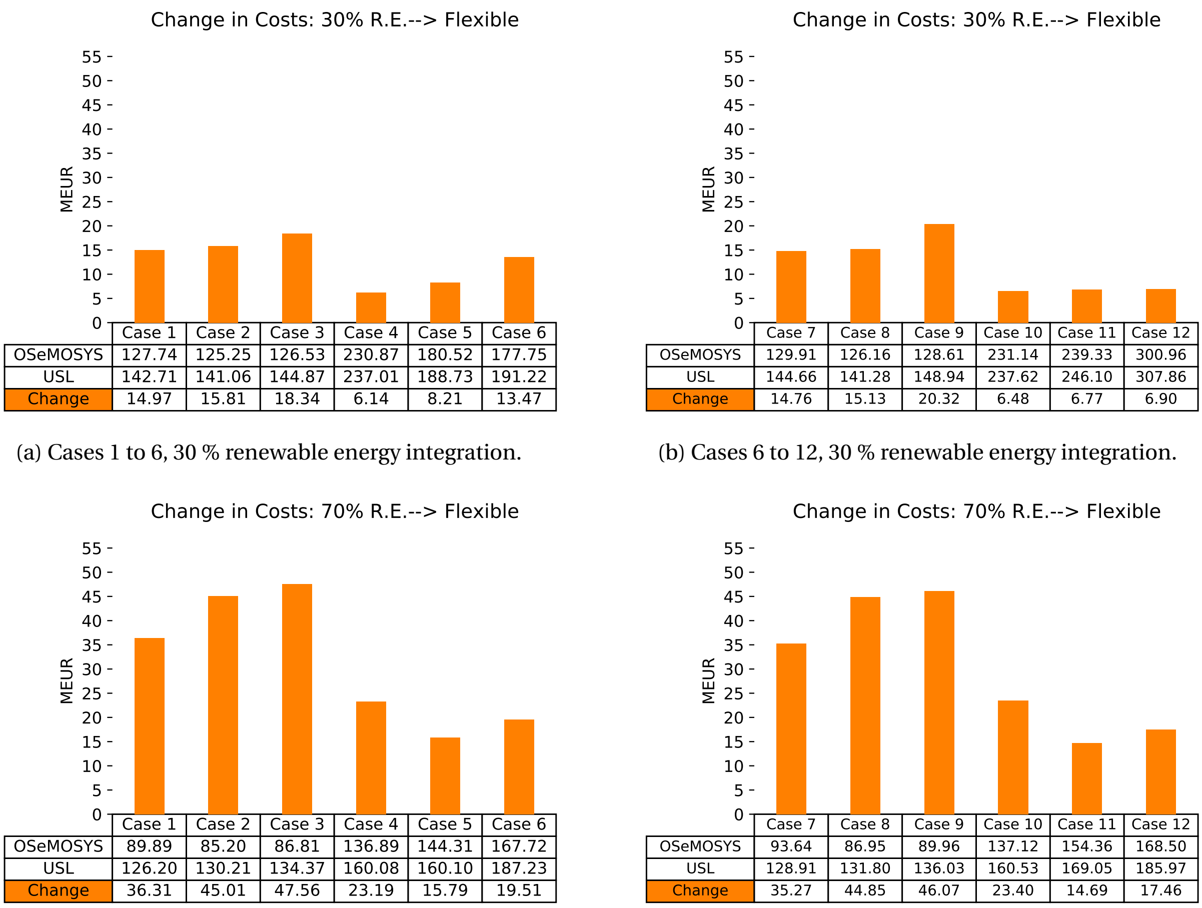
Cases	C^{acc} (kEUR/MW)			
	Nuclear	Coal	CCGT	COCGT
7	135.97	62.35	102.55	74.23
8	135.97	159.53	81.23	74.23
9	135.97	235.22	49.88	74.23
10	467.66	62.35	102.55	39.78
11	467.66	159.53	81.23	39.78
12	467.66	235.22	49.88	39.78

4.4.3. Unidirectional softlinking (USL): An equally insightful tool

For each of the twelve cases selected in Section 4.4, unidirectional softlinking (USL) as specified in Fig. 4.3 is applied to each of them. As previously explained in Subsection 4.3.3, it is enough for the USL to only consider the total variable cost of operation and the total start-up cost, since installed capacities do not change and thus resulting total fixed costs do not either. Furthermore, all cases are tested for 30 % and 70 % renewable energy (R.E.) integration in a system with power plants with flexible and inflexible parameters. Lastly, only the change in cost between OSeMOSYS and USL is plotted, as this is the determining variable to show the value of USL as a tool.

Fig. 4.7 shows the results for the flexible parameters. At 30 % R.E. integration, Case 4 has the lowest mismatch, 6.14 MEUR (see Fig. 4.7a); while Case 9 has the highest mismatch, 20.32 MEUR (see Fig. 4.7b). At 70 % R.E. integration, Case 11 has the lowest mismatch, 14.69 MEUR (see Fig. 4.7d); while Case 3 has the highest mismatch, 47.56 MEUR (see Fig. 4.7c). On the other hand, Fig. 4.8 shows the results for the inflexible parameters. At 30 % R.E. integration, Case 12 has the lowest mismatch, 12.71 MEUR (see Fig. 4.8b); while Case 9 has the highest mismatch, 42.34 MEUR (see Fig. 4.8b). At 70 % R.E. integration, Case 6 has the lowest mismatch, 18.53 MEUR (see Fig. 4.8c); while Case 8 has the highest mismatch, 52.11 MEUR (see Fig. 4.8d).

In summary, in a system with fixed parameters (either flexible or inflexible), a higher R.E. integration also means greater mismatches between OSeMOSYS and USL. In the specific scenario of a system with flexible parameters at either 30% or 70% R.E. integration, mismatches are between [6.14, 47.56] MEUR for all cases. Within the same considerations, mismatches in a system with inflexible parameters are between [12.71, 52.11] MEUR. This then suggests that inflexible parameters tend to increase the mismatches. As a matter of fact, it can be observed from the plots that for all individual cases, except for Case 6 at 70 % R.E. integration, inflexible parameters provoke higher mismatches in comparison to flexible parameters. This hence corroborates the results obtained in Subsection 4.3.3 for only two power plants. With these conclusions in mind, it can be stated that USL remains as a valuable tool to capture the several MEURs of underestimated costs by OSeMOSYS.



	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
OSeMOSYS	127.74	125.25	126.53	230.87	180.52	177.75
USL	142.71	141.06	144.87	237.01	188.73	191.22
Change	14.97	15.81	18.34	6.14	8.21	13.47

(a) Cases 1 to 6, 30 % renewable energy integration.

	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
OSeMOSYS	129.91	126.16	128.61	231.14	239.33	300.96
USL	144.66	141.28	148.94	237.62	246.10	307.86
Change	14.76	15.13	20.32	6.48	6.77	6.90

(b) Cases 6 to 12, 30 % renewable energy integration.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
OSeMOSYS	89.89	85.20	86.81	136.89	144.31	167.72
USL	126.20	130.21	134.37	160.08	160.10	187.23
Change	36.31	45.01	47.56	23.19	15.79	19.51

(c) Cases 1 to 6, 70 % renewable energy integration.

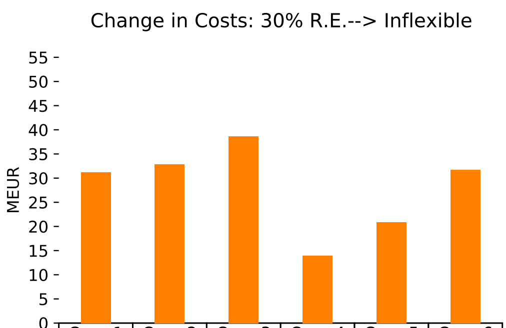
	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
OSeMOSYS	93.64	86.95	89.96	137.12	154.36	168.50
USL	128.91	131.80	136.03	160.53	169.05	185.97
Change	35.27	44.85	46.07	23.40	14.69	17.46

(d) Cases 6 to 12, 70 % renewable energy integration.

Figure 4.7: Change in costs from OSeMOSYS (only total variable cost of operation) to unidirectional softlinking (USL; actual total variable cost of operation and total start-up cost). The cases considered are as per Table 4.2 and Table 4.3 for power plants with a varying annualized capital cost C^{acc}. Power plants also have flexible parameters, while the system is subject to different levels of renewable energy integration.

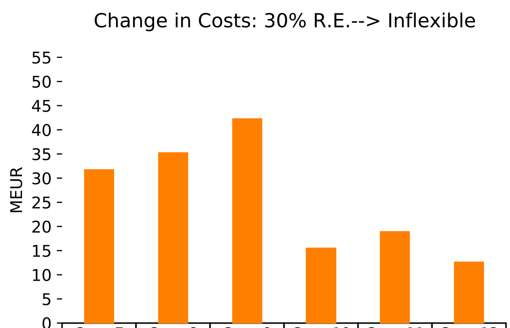
4.4.4. Bidirectional softlinking (BSL): Does a feedback loop scale?

As it was demonstrated in Subsection 4.4.3, the unidirectional softlinking (USL) method provides insightful results regarding missing costs for a four-generator system with different capital costs and discount rates. This hence validates the first experiments with only two power plants performed in Subsection 4.3.3. The next aspect to examine is whether the reduction in costs achieved by the bidirectional softlinking (BSL) method and the adjustments defined in Subsection 4.3.4 can also be scaled to the four-generator system with varying conditions. Here, only the *Variable cost re-scale* adjustment is considered, as this one was



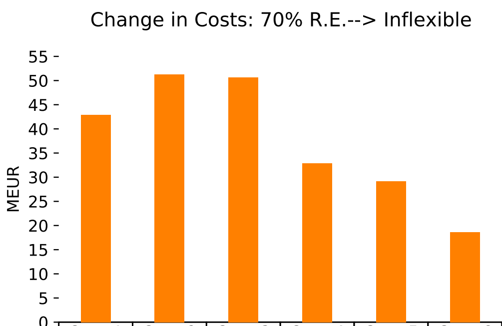
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
OSeMOSYS	127.74	125.25	126.53	230.87	180.52	177.75
USL	158.91	158.10	165.09	244.77	201.31	209.44
Change	31.17	32.85	38.56	13.90	20.79	31.69

(a) Cases 1 to 6, 30 % renewable energy integration.



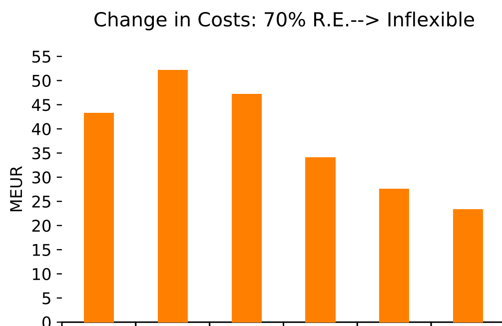
	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
OSeMOSYS	129.91	126.16	128.61	231.14	239.33	300.96
USL	161.67	161.44	170.96	246.73	258.32	313.66
Change	31.77	35.29	42.34	15.59	18.98	12.71

(b) Cases 6 to 12, 30 % renewable energy integration.



	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
OSeMOSYS	89.89	85.20	86.81	136.89	144.31	167.72
USL	132.72	136.43	137.42	169.77	173.47	186.25
Change	42.83	51.23	50.61	32.88	29.16	18.53

(c) Cases 1 to 6, 70 % renewable energy integration.



	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
OSeMOSYS	93.64	86.95	89.96	137.12	154.36	168.50
USL	136.90	139.06	137.13	171.23	181.90	191.80
Change	43.26	52.11	47.17	34.11	27.54	23.29

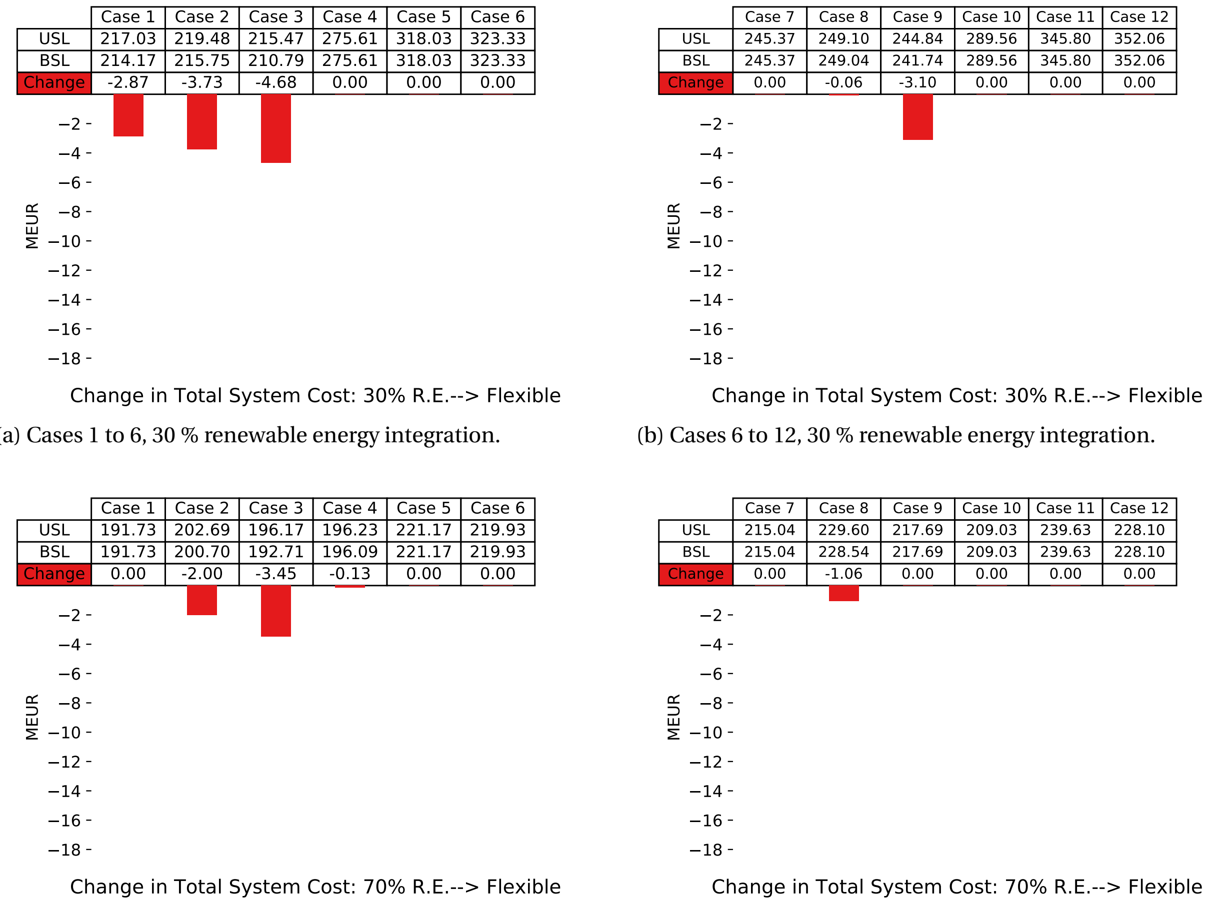
(d) Cases 6 to 12, 70 % renewable energy integration.

Figure 4.8: Change in costs from OSeMOSYS (only total variable cost of operation) to unidirectional softlinking (USL; actual total variable cost of operation and total start-up cost). The cases considered are as per Table 4.2 and Table 4.3 for power plants with a varying annualized capital cost C^{acc}. Power plants also have inflexible parameters, while the system is subject to different levels of renewable energy integration.

proven to need fewer iterations than the *Energy limit and SU costs* adjustment for reaching the lowest total system cost. For the BSL tests, the same conditions as for USL in Subsection 4.4.3 apply: four generators with flexible and inflexible parameters, n_{cases} = 12 cases, annualized capital costs as per Table 4.2 and Table 4.3, and 30 % and 70 % of R.E. integration.

In this context, Fig. 4.9 contains the results for the flexible parameters. At 30 % R.E. integration, the adjustment does not achieve a reduction for seven cases (see Fig. 4.9a and Fig. 4.9b). The lowest reduction is 0.06 MEUR in Case 8, while the highest one is 4.68 MEUR in Case 3. At 70 % R.E. integration, the adjustment does not achieve a reduction for eight cases (see Fig. 4.9c and Fig. 4.9d). The lowest reduction is 0.13 MEUR for Case 4, while the highest reduction is 3.45 MEUR for Case 3. Furthermore, the results for the inflexible parameters are shown in Fig. 4.10. At 30 % R.E. integration, the adjustment does not achieve a reduction for four cases (see Fig. 4.10a and Fig. 4.10b). The lowest reduction is 4.20 MEUR in Case 4, while the highest one is 16.30 MEUR in Case 3. Finally, at 70 % R.E. integration, the adjustment does not achieve a reduction for six cases (see Fig. 4.10c and Fig. 4.10d). The lowest reduction is 0.10 MEUR in Case 1, while the highest one is 8.26 MEUR in Case 7.

Taking these results into consideration, it can be concluded that the *Variable cost re-scale* adjustment has little effect on the four-generator system with flexible parameters. On the other hand, it does have a more pronounced influence on a system with inflexible parameters. However, it can be noted as well that the level of R.E. integration plays a significant role. For either flexible or inflexible parameters, the adjustment has a stronger effect at 30 % R.E. integration, while it gets much more reduced at 70 % R.E. integration. This in fact turns out to be counter-intuitive as the mismatches between OSeMOSYS and USL are greater at higher R.E. levels (as shown in Subsection 4.4.3) and should thus be reflected in further reductions in the total system cost after the BSL process. It can then be stated that the sizable effect of the *Variable cost re-scale* adjustment, as first suggested in Subsection 4.3.4 for a two-generator system, does not scale particularly well to the four-generator system with varying conditions.



(a) Cases 1 to 6, 30 % renewable energy integration.

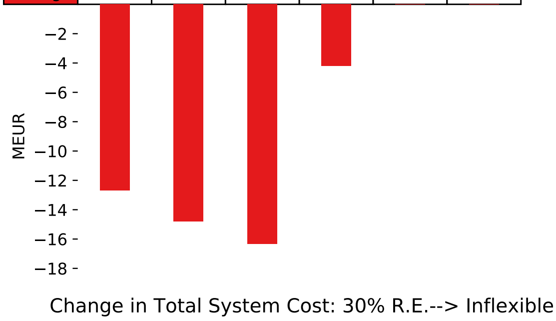
(b) Cases 6 to 12, 30 % renewable energy integration.

(c) Cases 1 to 6, 70 % renewable energy integration.

(d) Cases 6 to 12, 70 % renewable energy integration.

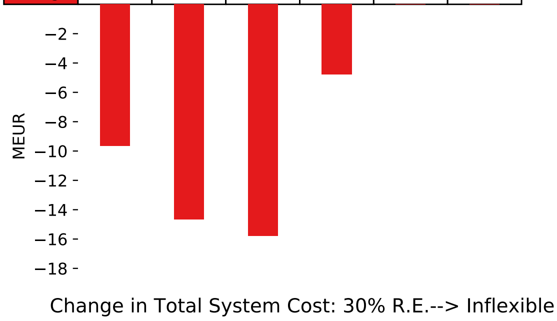
Figure 4.9: Change in costs from unidirectional softlinking (USL; initial total capital cost, actual total variable cost of operation and total start-up cost) to bidirectional softlinking (BSL; lowest resulting sum of total capital cost, actual total variable cost of operation and total start-up cost). The cases considered are as per Table 4.2 and Table 4.3 for power plants with a varying annualized capital cost C^{acc}. Power plants also have flexible parameters, while the system is subject to different levels of renewable energy integration. The adjustment considered is *Variable cost re-scale*.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
USL	233.23	236.52	235.69	283.38	330.61	341.55
BSL	220.55	221.73	219.38	279.17	330.61	341.55
Change	-12.69	-14.79	-16.30	-4.20	0.00	0.00



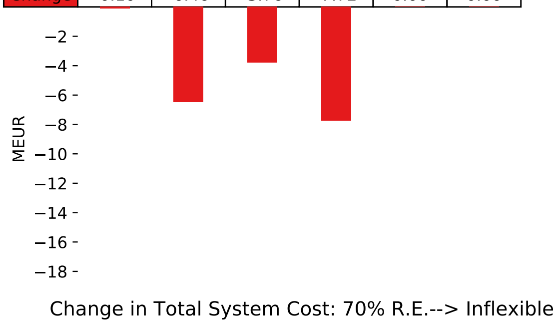
(a) Cases 1 to 6, 30 % renewable energy integration.

	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
USL	262.38	269.26	266.86	298.67	358.01	357.86
BSL	252.75	254.62	251.08	293.92	358.01	357.86
Change	-9.63	-14.64	-15.78	-4.75	0.00	0.00



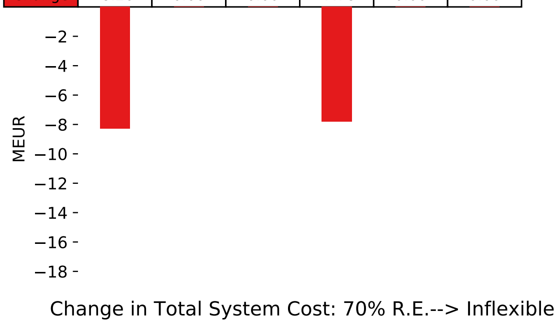
(b) Cases 6 to 12, 30 % renewable energy integration.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
USL	198.24	208.91	199.21	205.92	234.54	218.95
BSL	198.15	202.45	195.45	198.20	234.54	218.95
Change	-0.10	-6.46	-3.76	-7.72	0.00	0.00



(c) Cases 1 to 6, 70 % renewable energy integration.

	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
USL	223.03	236.86	218.78	219.74	252.48	233.93
BSL	214.77	236.86	218.78	211.95	252.48	233.93
Change	-8.26	0.00	0.00	-7.78	0.00	0.00



(d) Cases 6 to 12, 70 % renewable energy integration.

Figure 4.10: Change in costs from unidirectional softlinking (USL; initial total capital cost, actual total variable cost of operation and total start-up cost) to bidirectional softlinking (BSL; lowest resulting sum of total capital cost, actual total variable cost of operation and total start-up cost). The cases considered are as per Table 4.2 and Table 4.3 for power plants with a varying annualized capital cost C^{acc}. Power plants also have inflexible parameters, while the system is subject to different levels of renewable energy integration. The adjustment considered is *Variable cost re-scale*.

4.5. Case study III: Multiple generators, R.E. investment and storage arbitrage

4.5.1. Purpose and methodology

This case study is the third experiment that intends to answer the second specific research question in Section 1.6. The purpose of this case study is to understand through the use of bidirectional softlinking how long-term investments can be affected with two new circumstances in the power system: 1) that renewable energy is curtailed if not consumed and 2) that batteries in the system perform arbitrage for economic benefits. In contrast to Case study I in Section 4.3 and Case study II in Section 4.4, renewable energy sources are now an investment option, while the newly added storage devices actively react to electricity prices. Moreover, in order to recreate a more realistic scenario, the whole Dutch fleet of conventional generators is included in the optimization problem and their variable costs are used to form the electricity price in each time block. For the specific steps that have been used to build this scenario, please refer to Appendix A.3.

4.5.2. Description

In summary, Case study III considers investments in several types of power system technologies to supply the Dutch electric demand in 2030. With regards to conventional generation, investments in combined-cycle gas turbine (CCGT) and combustion-cycle gas turbine (COCGT) are considered. Nuclear is not an investment option, but the existing 485 MW Borselle nuclear power plant is still part of the mix. All existing coal power plants in the Dutch system are removed and coal is no longer available for investment either. Financial parameters for conventional generators are given as per Table 2.2. With regards to renewable energy sources, offshore wind, onshore wind and solar energy are regarded for investment. Financial parameters for renewable energy sources are given as per Table 4.4. Finally, storage units to be included are batteries, which remain too expensive to be deployed in the presence of conventional generation, so they are simply considered to already be a part of the system. On the other hand, these units may also be considered as private investments, which are then external to a central planner (see Subsection 2.2.3) and thus would not be a part of the investment optimization either.

On the side of power system operations, the amount of installed capacity per conventional generation technology is distributed along fifty power plants currently existing in the Netherlands. The list of power plants and their corresponding technology types are obtained from the tool set **powerplantmatching** [114]. The marginal costs of power plants from the same technology type are not necessarily the same (see Appendix A.3). Every power plant in the system bids according to its marginal cost, i.e. an energy-only market, described in Subsection 2.2.3. This thus forms the price of electricity. Renewable energy sources are dispatched at zero marginal cost, but they can be curtailed if their output is greater than demand. Storage units charge the electricity price for their services. They store energy when electricity prices are low and dispatch energy when they are high (i.e. arbitrage). There are a total of ten storage devices with an installed capacity within the range $i^{cap}_{bat} = [50, 150]$ MW. Their energy capacity e^{cap}_{bat} is equivalent to four hours of operation at full power, i.e. $e^{cap}_{bat} = [200, 600]$ MWh. The chosen values for the power capacity and energy capacity of batteries are in fact realistic. Recently, Southern California Edison announced seven battery projects with power capacity within the range $i^{cap}_{bat} = [50, 230]$ MW. The McCoy project with $i^{cap}_{bat} = 230$ MW will have an energy capacity of $e^{cap}_{bat} = 920$ MWh. The seven projects are expected to be completed by August, 2021 [115].

Table 4.4: Financial costs and technical parameters of three different technologies: large solar farm, onshore wind and offshore wind farms. Capital costs (exchange rate 1.333 USD per EUR) and operational lifetimes are obtained from Ref. [67]. The R=0.07 discount rate has been selected as it approximately corresponds to the market rate in deregulated or restructured markets [67]; which fits the European context.

Technology	Capital cost (kEUR/MW), C^{cc}	Operational lifetime (years), N	Annualized capital cost (kEUR/MW), C^{acc}, at discount rate, $R = 0.07$
Solar (PV)	770.01	25	66.08
Onshore wind (Won)	1625.80	25	139.51
Offshore wind (Woff)	2827.04	25	242.59

4.5.3. Bidirectional softlinking (BSL): A need?

As described in Subsection 4.5.2, the load profile is to be supplied- among others- by an array of technologies that is distributed over fifty conventional power plants. When in operation, these plants form the electricity price in an energy-only market. Batteries then offer their services at the electricity price, but their actions have a direct influence on the load profile. However, as stated, the load profile is one of the determining factors to choose the energy mix in the first place. This hence creates a cause-effect course that cannot be solved by a single stage optimization and requires instead an iterative process. The bidirectional softlinking (BSL) method consequently arises as a possible solution to the problem.

The approach here adopted for the BSL is specified in Fig. 4.12. Note that OSeMOSYS is used again as the ESOM, while PyPSA is now the UCED model (see Subsection 2.4.2 to find the reasons why PowerFactory has been replaced in this case). The methodology consists in tracking the total system cost, which is equal to the sum of total capital costs, fixed costs, actual variable cost of operation and curtailment costs. The adjustments injected in OSeMOSYS are two: 1) an upper bound on the actual energy consumed per renewable energy technology (the same as the one proposed in Subsection 4.3.4 for conventional generators) and 2) a new load profile. The new load profile sent back to OSeMOSYS contains the dispatch behavior of all batteries in the system.

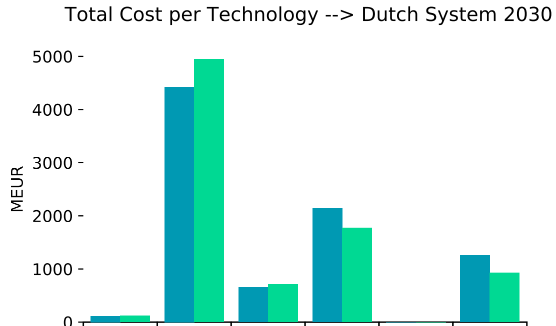
More specifically, consider the load point $P_b^{d,ose_{(k)}}$ at iteration k in OSeMOSYS. After the total battery dispatch per point $\sum_{bat} p_{bat,b}^{pyp_{(k)}}$ has been calculated in PyPSA, the load point $P_b^{d,ose_{(k+1)}}$ to be inserted in OSeMOSYS at iteration $k+1$ is:

$$P_b^{d,ose_{(k+1)}} = P_b^{d,ose_{(k)}} - \sum_{bat} p_{bat,b}^{pyp_{(k)}} \tag{4.9a}$$

$$b \subset T^{ose}. \tag{4.9b}$$

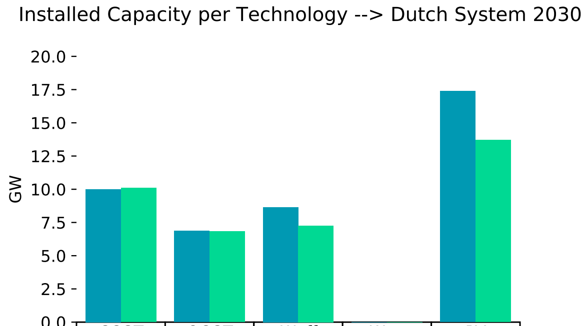
The results are shown in Fig. 4.11. Fig. 4.11a in specific shows the changes in total cost per technology. For the conventional generators (i.e. Nuclear, OCGT and COCGT) the total cost includes capital costs, fixed costs and variable cost of operation. On the other hand, the total cost for renewable energy sources is composed of capital costs and curtailment costs. As it is observed in the figure, nuclear has an increase of 4.23 MEUR after BSL is applied. This is only due to its variable cost of operation, as its installed capacity is not subject to change. Moreover, the total costs of OCGT and COCGT also increase by 517.82 MEUR and 54.36 MEUR, respectively. For offshore wind and solar energy, the total cost decreases by 368.24 MEUR and 329.19 MEUR, respectively. In summary, the net change in total cost after BSL is $\Delta c_{sys}^{total} = -121.02$ MEUR (i.e. savings).

By observing the changes in installed capacities in Fig. 4.11b, it is possible to dissect the total cost per technology. For instance, the change in total cost for PV can be expressed in terms of the change in its installed capacity and the change in the total cost to curtail its energy: $\Delta c_{pv}^{total} = C_{pv}^{acc} \Delta i_{pv}^{cap} + \Delta c_{pv}^{curt} =$ 329.19 MEUR. Solving for the change in curtailment costs then gives $\Delta c_{pv}^{curt} = 85.35$ MEUR. This indicates that OSeMOSYS can severely misrepresent curtailment costs if short-term power system operations are not included. Finally, while OCGT and COCGT have small changes in their installed capacities, the ones in offshore wind and PV are both noteworthy: Δi^{cap}_{woff} = 1.39 GW and Δi^{cap}_{pv} = 3.69 GW, respectively. Their decrease after BSL is significant, as the vast of curtailed energy becomes evident when short-term power system operations are taken into consideration.



	Nucelar	CCGT	OCGT	Woff	Won	PV
Before BSL	115.45	4428.19	654.90	2141.23	0.00	1256.62
After BSL	119.68	4946.02	709.26	1772.99	0.00	927.43
Change	4.23	517.82	54.36	-368.24	0.00	-329.19

(a) Total cost per technology. For conventional sources, the total cost consists of the capital costs, fixed costs and variable cost of operation. For renewable energy sources, the total cost is: capital costs and curtailment costs.



	CCGT	OCGT	Woff	Won	PV
Before BSL	10.01	6.86	8.66	0.00	17.39
After BSL	10.10	6.83	7.27	0.00	13.70
Change	0.09	-0.03	-1.39	0.00	-3.69

(b) Installed capacity per technology. By knowing the changes in their installed capacities, the changes in the other costs can also be obtained. The greatest changes in installed capacities occur in offshore wind energy (i.e. Δi^{cap}_{woff} = 1.39 GW) and solar energy (i.e. Δi^{cap}_{pv} = 3.69 GW).

Figure 4.11: Comparison between the total cost and installed capacity per technology, before and after the application of bidirectional softlinking (BSL).

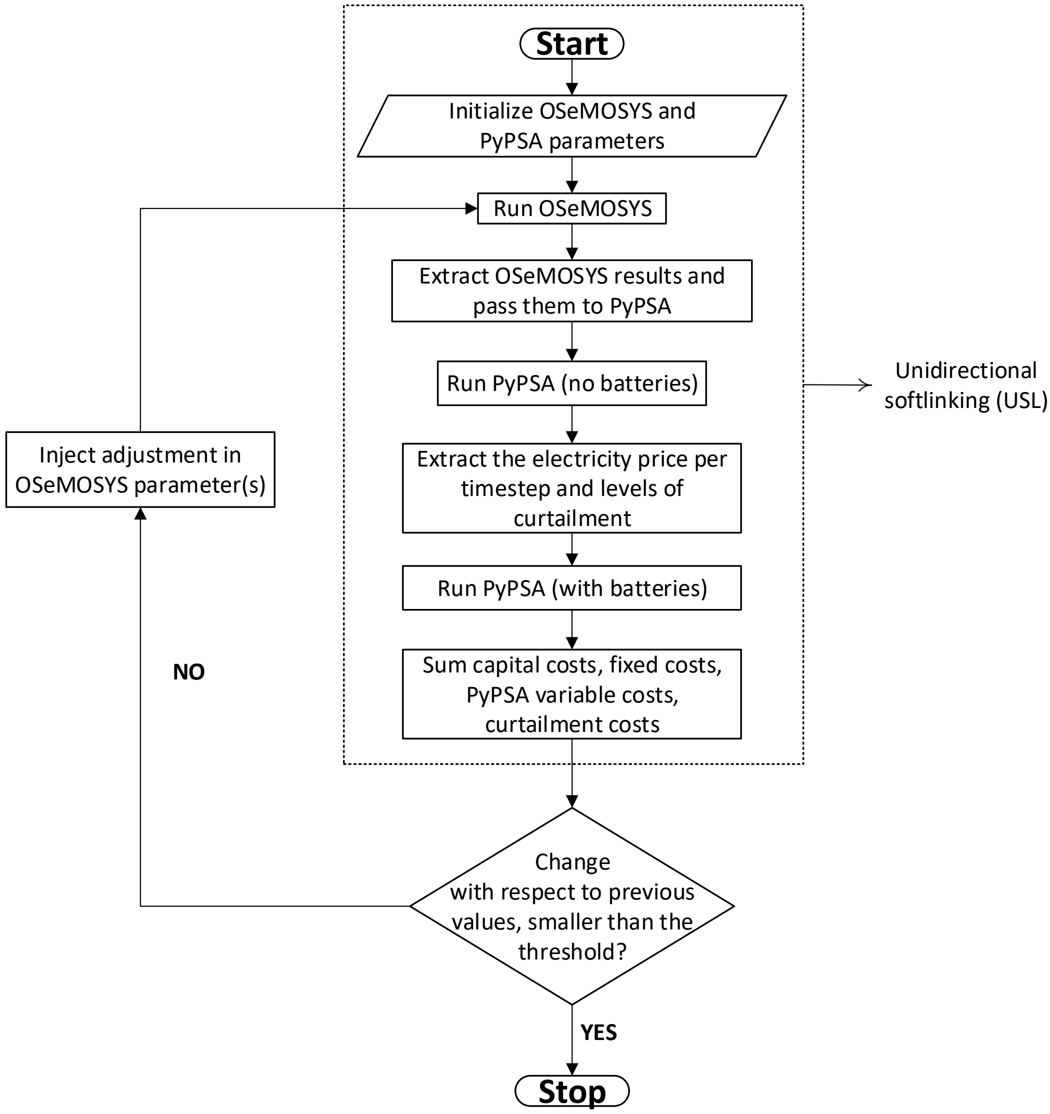


Figure 4.12: Flowchart of the bidirectional softlinking (BSL) coupling method between OSeMOSYS and PyPSA. The process consists in an unidirectional softlinking (USL; a subset of BSL as illustrated by the dashed block), whose results are fed back to OSeMOSYS in each iteration via designed adjustments. The value to be tracked in this specific BSL is the total system cost, which is composed of capital costs, fixed costs and variable cost of operation for conventional technologies and curtailment costs for renewable energy sources.

5
Concluding remarks

5.1. Introduction

This chapter concludes this thesis. More specifically, it consolidates the several aspects covered in Chapter 2, Chapter 3 and Chapter 4 in order to provide comprehensive answers to the three specific research questions introduced in Chapter 1. Furthermore, this chapter proposes future lines of work that can be followed, as well as new methodologies that can be adopted to capture the influence of short-term operations in long-term investments.

5.2. Conclusions and discussions through the specific research questions

As indicated in Chapter 1, the urgent energy transition to reduce environmental deterioration has stimulated a wave of investment needs in a new generation of technologies. For example, these technologies include electric vehicles, distributed generation, demand side management, etc. The *short-term* actions of these technologies, however, induce unprecedented effects on the power system (e.g. less predictable load patterns). This suggests that assumptions and techniques used in the past to decide upon long-term investments in the power system must be reconsidered. This idea is encompassed in a single central question in the same chapter: *How can short-term operations influence long-term investments in future electric power systems?* Subsequently, Chapter 1 covers the definitions, timescales and stakeholders around this question and proceeds to define the following specific research questions:

1. How can long-term investments and short-term operations in the power system be modeled and exchange *relevant* information?

   - **Conclusion:** As discussed in Chapter 2, energy system models are typically used to represent long-term investments in interconnected energy sectors, including the power system. In this thesis, a specific type of model is selected: energy system *optimization* model (ESOM), as it provides a rich technological representation and a clear cost evolution over time. On the other hand, short-term power system operations are represented by a unit commitment and economic dispatch (UCED) model, which provides a set equations for system costs, generator and grid dynamics, levels of curtailed renewable energy, etc.

     In Chapter 3, three different ways for ESOMs and UCED models to exchange information are reviewed. The first one is *co-optimization* in which both models are integrated as a single optimization problem. The other two are *unidirectional softlinking* (USL) and *bidirectional softlinking* (BSL). In this thesis, these two approaches are limited to inserting the variable results from one of the models as parameters into the other model. Among co-optimization, USL and BSL, the latter is selected as the best option due to its flexibility to be expanded in terms of equations and other energy sectors to be considered beside the power system itself.

The *relevant* information to be exchanged between an ESOM and a UCED model truly depends on the application and the answers being sought. As it was observed in Chapter 3 through the literature review process (particularly co-optimization and BSL), some studies add the UCED constraints as a way to *inform* the ESOM about certain flexibility requirements. Other studies are also interested in a similar information, but include storage units and other sources of flexibility to unveil their potential. Other authors are interested instead in the revenue of power producers (e.g. conventional generators or hydro-power) to enable changes in the energy mix. The *relevant* information to be exchanged in this thesis is further explained in the next two specific research questions.

- **Discussion:** ESOMs are mostly used to model several years, regions and technologies so they are employed with a very low-time resolution and a lack of short-term operations equations. The most significant disadvantage of a low-time resolution is that detailed information (e.g. renewable power fluctuations) may be lost. However, Chapter 4 elaborates on a data clustering algorithm to refine the information to be inserted in the ESOM, while keeping the same low-time resolution. The algorithm shows that, for instance, the system load duration curve can be reconstructed to an acceptable degree.

  In the same context, Chapter 2 suggests that a low-time resolution and a lack of short-term operations equations can be both improved at once by coupling ESOMs with UCED models. UCED models are not only capable of modeling short-term power system operations, as aforementioned, but also provide a finer resolution that keeps the chronology of events and a set of more accurate equations related to system costs and dynamics. In this thesis, OSeMOSYS-PuLP is selected as the ESOM and PowerFactory 2019 and PyPSA as the UCED models.

  Among the three ways to couple ESOMs and UCED models in this thesis (see Chapter 3), unidirectional softlinking (USL) is to be considered as a valuable tool to obtain a first impression of the costs that are misrepresented by the ESOM in comparison to the UCED model. On the other hand, co-optimization and bidirectional softlinking (BSL) should be both regarded as coupling methods in which both models exert a certain influence on each other. Chapter 3 advocates for BSL as the best option among the two, since it possibly allows for a greater flexibility. This means that each model can be expanded (e.g. more equations, inclusion of other systems) based on their own nature, as they remain as separate entities throughout the simulations. Nonetheless, BSL itself carries significant challenges regarding the tractability of the simulations required, and the selection of the convergence criteria and the information to be exchanged.

2. Under what circumstances do short-term operations become influential on long-term investments in the power system?

   - **Conclusion:** After the selection of bidirectional softlinking (BSL) as the coupling method in this thesis, three case studies are formulated in Chapter 4. The intention of these case studies is twofold: 1) test the applicability of BSL and 2) understand via BSL, under what circumstances short-term operations are capable of influencing long-term investments in the power system.

     For Case study I (see Section 4.3) and Case study II (see Section 4.4), the unidirectional softlinking was first applied to obtain an impression of how significantly OSeMOSYS misrepresents the total system cost due to its low-time resolution and lack of short-term operations equations. Two important observations from these case studies are: 1) the misrepresented costs become higher with a greater level of renewable energy integration and 2) the misrepresented costs become higher when power plants in the energy mix are less flexible. Therefore, this implies that the expected greatest mismatch occurs under two specific conditions: an inflexible system with a high R.E. integration.

     Similarly, it would be expected that it is precisely under these two circumstances that short-term operations become more influential on long-term investments in the power system. However, as it was demonstrated in Case study II, this highly depends on the adjustment that is sent back to OSeMOSYS. The adjustment here selected for further study, i.e. *Variable cost re-scale*, provided in fact counter-intuitive results to some extent and did not scale well enough to produce the same successful results as in Case study I. This suggests that other adjustment designs may be better at exploiting the information first obtained by the USL when there are more than two generators.

     For Case study III (see Section 4.5), the motivation is different. There, the idea is to create two specific circumstances: 1) that renewable energy (R.E.) is curtailed if not consumed and 2) that batteries in the system perform arbitrage for economic benefits. These two conditions prove to have a relevant influence over the long-term investment decisions taken by OSeMOSYS. In comparison to Case study I and Case study II, the adjustments sent back to OSeMOSYS in Case study III (i.e. battery behavior injected in the demand curve and limits on renewable energy output) seem to be more intuitive and direct, and thus are considered to better represent short-term operations on the long run.

   - **Discussion:** In Case study I (see Section 4.3) and Case Study II (see Section 4.4), it can be observed that an inflexible system with a high R.E. integration generates the highest mismatch between OSeMOSYS and PowerFactory. In fact, the greatest cost mismatch occurs between the variable costs of operation, since start-up costs are only a small percentage of the total mismatch. This can be explained by revisiting the two-generator system in Case study I. There, the peak demand is $P^{d,peak}$ = 1000 MW and the installed capacities for nuclear and coal are: i^{cap}_{nuc} = 243.94 MW, i^{cap}_{coal} = 687.55 MW, respectively. In the UCED model, there are specific points in time in which renewable power is high enough to drop the net demand such that

$i_{nuc}^{cap} \leq P_b^d \leq P_{coal}^{min} \cdot i_{coal}^{cap}$. As the net demand still remains above what the nuclear power plant can supply, the coal power plant remains on and it is set to $p_{coal,b}^{pf} = P_{coal}^{min} \cdot i_{coal}^{cap}$. In contrast, OSeMOSYS, which is free of inter-temporal constraints, calculates $p_{coal,b}^{ose} = P_b^d - i_{nuc}^{cap}$. This creates a difference in calculated power, $\Delta p_{coal,b} = p_{coal,b}^{pf} - p_{coal,b}^{ose} > 0$, followed by a mismatch in the variable cost of operation, $\Delta c_{coal,b}^{var} = C_{coal}^{var} \cdot \Delta p_{coal,b}$. This behavior repeated throughout every time block, b, in which the same conditions apply, results in a greater overall mismatch. Finally, the greater P_{coal}^{min} is (i.e. a less flexible system) and the more often the net load P_b^d drops (i.e. higher R.E. integration), the bigger and more recurrent the mismatches are per time block. Of course, UCED is in fact a full optimization problem and several other factors can influence this result, but the aforementioned explanation constitutes a simple approach to understand the reason why mismatches are greater in an inflexible system with a higher R.E. integration.

In these two case studies, the *Variable cost re-scale* adjustment is assessed. In Case study I, this adjustment quickly and accurately pushes the system to its optimal value (i.e. as corroborated by a space sweep). However, when two more generators are added in Case study II, the adjustment seems to have a reduced influence, at least when compared to the scale of the mismatches between OSeMOSYS and PowerFactory. Furthermore, the adjustment has a greater effect on the inflexible system, which is consistent with the USL result (i.e. an inflexible system has a greater cost mismatch). Nonetheless, it produces an opposite outcome with respect to the level of R.E. integration, as the adjustment is more influential when the share of R.E. is lower. This is counter-intuitive, as aforementioned, since it contradicts the USL result in which cost mismatches are greater at a higher R.E. integration level. Hence, it can be concluded that the *Variable cost re-scale* adjustment does not posses the best design to capture the influence of system flexibility and renewable energy together.

Finally, Case study III (see Section 4.5) is particularly interesting as it tackles the problem from the perspective of a central planner that must deal with other existing components in the network. This case study shows that the optimal energy mix (i.e. conventional and renewable energy sources) highly depends on how, for instance, batteries react to electricity prices or how much R.E. energy is to be curtailed. Most importantly, applying the BSL iterative process allows a central planner to save $\Delta c_{sys}^{total} = -121.02$ MEUR for this specific case. Methodologically, this case study suggests that some feedback loops, i.e. adjustments, in the BSL can be inserted more naturally. In this way, a modeler does not need to implement a sophisticated design to capture extra costs and constraints, as done for Case study I and Case study II.

3. How impactful are the changes produced by short-term operations in comparison to the scale of long-term investments in the power system?

   - **Conclusion:** In Case study I (see Section 4.3) and Case Study II (see Section 4.4), the true long-term investment costs are the ones calculated by the unidirectional softlinking (USL): OSeMOSYS calculates the capital costs of the selected energy mix and PowerFactory calculates the total start-up cost and variable cost of operation. On the other hand, the influence of the short-term power system operations on the long-term investments is measured according to the cost reduction that the bidirectional softlinking (BSL) induces on the final total system cost. In Case study I, the reduction effects are visible: 4.18% (from 360.04 MEUR to 345 MEUR) and a 6.41% (from 376.11 MEUR to 352 MEUR) for a flexible and an inflexible system, respectively. In Case II, the greatest relative reduction is 6.92% (from 235.69 MEUR to 219.38 MEUR) for Case 3 at 30 % R.E. integration and inflexible parameters. However, there are several cases in which very little to no improvement is obtained. In Case study III, the relative improvement is of only 1.41% (8596.39 MEUR of total system cost and a reduction of $\Delta c_{sys}^{total} = -121.02$ MEUR). Nevertheless, it is to be considered that unit commitment constraints are not a part of the set-up due to the number of power plants in the system (see Appendix A.3 for further details). This would add extra costs, but potentially also greater relative savings, as it does for the other two case studies.

Finally, the *impact* of short-term operations can also be assessed in terms of the changes in installed capacity per technology from the resulting energy mix. For example, the changes in installed capacity on Case study III are: 19.52% in offshore wind reduction (from 8.66 GW to 7.27 GW) and 21.22% in solar reduction (from 17.39 GW to 13.70 GW). This suggests that the inclusion of curtailment costs at every timestep can significantly change the final mix, even in the presence of batteries that are to store much of the overproduction of renewable energy.

- **Discussion:** The *impact* of short-term operations on the long-term investments can probably be measured in different ways, depending on what the modeler intends to improve in the first place and how this leads to changes in the long run. In this thesis, it is proposed that the influence of short-term operations should be measured in terms of changes in installed capacities and most importantly in terms of the reduction of the overall system cost. For Case study I, the percentages mentioned above about cost improvements are truly conclusive, as corroborated by space weep (see Subsection 4.3.5). Nonetheless, this does not entirely hold for Case study II. First, this is because the adjustment seems to produce irrelevant changes in comparison to the big mismatches calculated by the unidirectional softlinking (USL); and second, because there is no simple way to prove that the adjustment pushes the system to its best possible outcome (i.e. space sweep is too computational expensive with four generators, and linearity cannot be maintained in order to obtain the true optimal). Finally, Case study III is more directed towards a methodological development, rather than achieving the lowest possible costs. Even then, placing the total system cost as the central variable to track, allows to obtain important overall reductions, as aforementioned.

5.3. Recommendations and future work

This thesis elaborates on the concept of *bidirectional softlinking* (BSL) between ESOMs and UCED models, the advantages that the method provides to power system planning (e.g. reductions in total system costs) and the remarkable challenges that must be faced to enable its application. Throughout this thesis, the author has undergone a very steep learning curve, and he would like to share his knowledge in terms of recommendations and future research directions that might help someone entering the field have a clearer idea and a more solid base to start:

- Test other variations of bidirectional softlinking (BSL).

  In this thesis, BSL has been limited to only applying a specific type of bidirectional softlinking that consists in injecting variable results as parameters (see Chapter 3). This may not be the most convenient way, as not all variable results naturally fit some parameters. For example, here the start-up cost has been injected as a fixed cost in OSeMOSYS in the *Energy limit and SU costs* adjustment. However, the start-up cost represents a different type of dynamics that is not entirely related to the size of the power plant (i.e. as the fixed cost is), but also to the behavior of other plants in the system. Therefore, a better way may be to use a heuristic formulation to set up the start-up cost as a parameter in OSeMOSYS in such a way that it can be refined by PowerFactory results.

  However, aggregating parameters in this way may not be the best possible option on the long run either. This is because a new formulation would be needed every time a new parameter is to be adjusted or refined in OSeMOSYS. On the other hand, it may be possible to only add new formulations for the most significant parameters (e.g. as it was shown here that mismatches in variable costs are much more significant than start-up costs) to circumvent this problem.

- Create better adjustment designs.

  The idea of keeping both models separate while implementing as few changes as possible to their inherent structure is still the preferred concept. This is particularly important to avoid interfering with the modeler's work and instead establish BSL as a communication mechanism between models. In order to achieve this idea, adjustments are to be the de facto enabler of the information exchange process, but their designs need to be more elaborated, precise, and remain non-intrusive. An elaborated adjustment design needs to have a mathematical justification, perhaps even convey an idea on how such adjustment is supposed to push the investment model to lower total system costs. This kind of design may be achieved by doing further research in the coupling of optimization problems, which would help provide a basis on how softwares can also be linked.

- Implement bidirectional softlinking (BSL) to a multi-year investment horizon.

  After the methodology to be followed has been well defined by testing the ESOM-UCED model coupling in just a single optimization period, a multi-year coupling can be implemented to obtain further insights into the potential of the BSL method. A multi-year investment horizon would eventually scale up to the timescale of a full generation expansion planning, including asset depreciation over time.

- Consider a case study in which there are no conventional power plants.

  Some visions predict a future without any conventional generators (e.g. Dutch Climate Agreement [8]). Therefore, it would be worth testing the BSL potential in a system with only renewable energy sources, storage units and sector coupling. Regarding the latter, other models can be used (e.g. an actual software to model gas networks).

Bibliography

[1] J. Olivier and J. Peters. Trends in global CO2 and total greenhouse gas emissions. Report summary, PBL Netherlands Environmental Assessment Agency, 2019.

[2] International Energy Agency. Global energy and CO2 status report: The latest trends in energy and emissions in 2018. Report, International Energy Agency, 2019.

[3] World Economic Forum. The future of electricity: New technologies transforming the grid edge. Report, World Economic Forum, 2017.

[4] M. Welsch, M. Howells, M. R. Hesamzadeh, B. Ó Gallachóir, P. Deane, N. Strachan, M. Bazilian, D. M. Kammen, G. Strbac L. Jones, and Holger Rogner. Supporting security and adequacy in future energy systems: The need to enhance long-term energy system models to better treat issues related to variability. *International Journal of Energy Research*, 39:377–396, 2015.

[5] K. Engeland, M. Borga, J. Creutin, B. Francois, Maria-Helena Ramos, and Jean-Philippe Vidal. Space-time variability of climate variables and intermittent renewable electricity production – a review. *Renewable and Sustainable Energy Reviews, Elsevier*, 79:600–617, 2017.

[6] A. van Stiphout. *Short-Term Operational Flexibility in Long-Term Generation Expansion Planning*. PhD thesis, KU Leuven, Leuven, Belgium, 1 2017.

[7] Y. Chai, Y. Xiang, J. Liu, C. Gu, W. Zhang, and W. Xu. Incentive-based demand response model for maximizing benefits of electricity retailers. *Journal of Modern Power Systems and Clean Energy, IEEE*, 7:1644 – 1650, 2019.

[8] The Goverment of the Netherlands. Climate agreement. Technical report, The Government of the Netherlands, 2019.

[9] International Renewable Energy Agency. Planning for the renewable future long-term modelling and tools to expand variable renewable power in emerging economies. Report, International Renewable Energy Agency, 2017.

[10] A. Schroder, F. Kunz, J. Meiss, R. Mendelevitch, and C. von Hirschhausen. Current and prospective costs of electricity generation until 2050. Reasearch report, Deutsches Institut für Wirtschaftsforschung, 2013.

[11] H. Bjornebye. *Investing in EU Energy Security: Exploring the Regulatory Approach to tomorrow's electricty production*. Kluwer Law International B.V., The Netherlands, 2010.

[12] World Economic Forum. The future of electricity attracting investment to build tomorrow's electricity sector. Report, World Economic Forum, 2015.

[13] European Commission. Study on regulatory incentives for investments in electricity and gas infrastructure projects. Report, European Commission, 2014.

[14] Fraunhofer ISI and Ecofys. The impact of risks in renewable energy investments and the role of smart policies. Report, Fraunhofer ISI and Ecofys, 2016.

[15] International Energy Agency. World energy investment 2019. Report, International Energy Agency, 2019.

[16] European Commission. Energy roadmap 2050. Report, European Commission, 2012.

[17] European Energy Industry Investments. Study for the European's parliament's committee on industry, Research and Energy (ITRE). Report, European Energy Industry Investments, 2017.

[18] International Energy Agency. Technology roadmap-energy storage, technical report. Report, International Energy Agency, 2014.

[19] P. Sauer. Time-scale features and their applications in electric power system dynamic modeling and analysis. San Francisco, CA, USA, 2011. Proceedings of the 2011 American Control Conference, IEEE.

[20] International Renewable Energy Agency. Power system flexibility for the energy transition (Part 1: Overview for policy makers outlines a range of options to scale up system flexibility). Report, International Renewable Energy Agency, 2018.

[21] International Renewable Energy Agency. Electricity storage valuation framework: Assessing system value and ensuring project viability. Report, International Renewable Energy Agency, 2020.

[22] S. Salm. *Investor-specific risk perception and return expectation for renewable energy projects: Evidence from the German and Swiss energy market.* PhD thesis, University of St. Gallen, St. Gallen, Switzerland, 5 2017.

[23] OECD. Mapping channels to mobilise institutional investment in sustainable energy. In *Green Finance and Investment*. OECD publishing, Paris, 2015.

[24] A. Prontera. Energy policy: Concepts, actors, instruments and recent developments. *World Political Science Review*, 5, 2009.

[25] Nordic Energy Regulators. A common definition of the system operators' core activities. Report, NordREG, 2017.

[26] L. Bird, J. Cochran, and Xi Wang. Curtailment: Experience and practices in the United States. Technical report, National Renewable Energy Laboratory, 2014.

[27] A. Knaut and S. Paulus. Hourly price elasticity pattern of electricity demand in the German day-ahead market. *EWI Working Paper, No. 16/07, Institute of Energy Economics at the University of Cologne (EWI), Köln*, 2016.

[28] B. S. Palmintier. *Incorporating operational flexibility into electric generation planning: Impacts and method for system design and policy analysis.* PhD thesis, MIT, Cambridge, MA, USA, 10 2012.

[29] M. Welsch, P. Dean, M. Howells, B. Ó Gallachóir, F. Rogan, M. Bazilian, and H. Rogner. Incorporating flexibility requirements into long-term energy system models – a case study on high levels of renewable electricity penetration in ireland. *Applied Energy, Elsevier*, 135:600–615, 2014.

[30] T. Burandta, B. Xiong, K. Löfflerb, and P. Oeib. Decarbonizing China's energy system – modeling the transformation of the electricity, transportation, heat, and industrial sectors. *Applied Energy, Elsevier*, 155, 2019.

[31] Williams J. H, B. Haley, F. Kahrl, J. Moore, A.D. Jones, M.S. Torn, and H. McJeon. Pathways to deep decarbonization in the United States. Technical report, The U.S. report of the Deep Decarbonization Pathways Project of the Sustainable Development Solutions Network and the Institute for Sustainable Development and International Relations, 2014.

[32] T. Kober, J. Falzon, B. van der Zwaan, K. Calvin, A. Kanudia, A. Kitous, and M. Labriet. A multi-model study of energy supply investments in Latin America under climate control policy. *Energy Economics, Elsevier*, 56:543–551, 2016.

[33] B. van der Zwaana, T. Kober, F. D. Longa, A. van der Laane, and G. J. Kramerg. An integrated assessment of pathways for low-carbon development in Africa. *Energy Policy, Elsevier*, 117:387–395, 2018.

[34] K. Loffler, T. Burandt, K. Hainsch, and P. Oei. Modeling the low-carbon transition of the European energy system - a quantitative assessment of the stranded assets problem. *Energy Strategy Reviews, Elsevier*, 26, 2019.

[35] K. Palmer-Wilson, J. Donald, B. Robertson, B. Lyseng, V. Keller, M. Fowle, C. Wade, S. Scholtysik, P. Wild, and A. Rowe. Impact of land requirements on electricity system decarbonisation pathways. *Energy Policy, Elsevier*, 129:193–205, 2019.

[36] J. Anjo, D. Neves, C. Silva, A. Shivakumar, and M. Howells. Modeling the long-term impact of demand response in energy planning: The Portuguese electric system case study. *Energy, Elsevier*, 165:456–468, 2018.

[37] F. Urban. *Low carbon transitions for developing countries.* Routledge, London, UK, 2014.

[38] H. Farzaneh. *Energy system modeling: principles and applications.* Springer, Berlin, Germany, 2019.

[39] A. Subramanian, T. Gundersen, and T. Adams II. Modeling and simulation of energy systems: A review. *Processes, MDPI*, 6, 2018.

[40] F. Mougouei and M. Mortazavi. Effective approaches to energy planning and classification of energy systems models. *International Journal of Energy Economics and Policy*, 7, 2017.

[41] L. Hall and A. Buckley. A review of energy systems models in the UK: Prevalent usage and categorisation. *Applied Energy, Elsevier*, 169:607–628, 2016.

[42] H. Ringkjob, P. Haugan, and I. Solbrekke. A review of modelling tools for energy and electricity systems with large shares of variable renewables. *Renewable and Sustainable Energy Reviews, Elsevier*, 96:440–459, 2018.

[43] J. Després, N. Hadjsaid, P. Criqui, and I. Noirot. Modelling the impacts of variable renewable sources on the power sector: Reconsidering the typology of energy modelling tools. *Energy, Elsevier*, 80:486–495, 2015.

[44] N. van Beeck. Classification of energy models. Master's thesis, Tilburg University and Eindhoven University of Technology, Tilburg and Eindhoven, The Netherlands, 5 1999.

[45] S. Jebaraj and S. Iniyan. A review of energy models. *Renewable and Sustainable Energy Reviews, Elsevier*, 10:281–311, 2006.

[46] S.C. Bhattacharyya SC and G.R. Timilsina. A review of energy system models. *International Journal of Energy Sector Management*, 4, 2010.

[47] S. Collins, J. P. Deane, K. Poncelet, E. Panos, R. C. Pietzcker, E. Delarue, and D. P. Ó Gallachóir. Integrating short term variations of the power system into integrated energy system models: A methodological review. *Renewable and Sustainable Energy Reviews, Elsevier*, 76:839–856, 2017.

[48] H. Shavandi, J. D. Fuller, and M. Pirnia. An efficient data driven model for generation expansion planning with operational constraints. 10 2018.

[49] M. Petitet. *Long-term dynamics of investment decisions in electricity markets with variable renewables development and adequacy objectives.* PhD thesis, Dauphine Uversité Paris and PSL, Paris, France, 11 2017.

[50] F. Olsina. *Long-term dynamics of liberalized electricity markets.* PhD thesis, Universidad Nacional de San Juan, San Juan, Argentina, 7 2005.

[51] A. Conejo, L. Baringo, S. J. Kazempour, and A. S. Siddiqui. *Investment in electricity generation and transmission: Decision making under uncertainty.* Springer International Publishing, Berlin, Germany, 2016.

[52] O. M. Babatunde, J. L. Munda, and Y. Hamam. A comprehensive state-of-the-art survey on power generation expansion planning with intermittent renewable energy source and energy storage. *International Journal of Energy Research, Wiley*, 43, 2019.

[53] A. Geiger. *Strategic plan investment planning under fuel and carbon price uncertainty.* PhD thesis, Karlsruhe Insitute of Technology, Karlsruhe, Germany, 4 2010.

[54] European Wind Energy Association. Powering Europe: Wind energy and the electrical grid. Report, European Wind Energy Association, 2010.

[55] C. Kemfert, D. Barbua, and V. Kalashnikov. Economic effects of the liberalization of the European electricity market- simulation results of a game theoretic modelling concept. 1 2003.

[56] T. Zhang, R. Baldick, and T. Deetjen. Optimized generation capacity expansion using a further improved screening curve method. *Electric Power Systems Research, Elsevier*, 124:47–54, 2020.

[57] T. Zhang and R. Baldick. Consideration of ancillary services in screening curve method. Denver, CO, USA, 2015. 2015 IEEE Power & Energy Society General Meeting, IEEE.

[58] T. Zhang and R. Baldick. Consideration of existing capacity in screening curve method. *IEEE Transactions on Power Systems , IEEE*, 32:3038 – 3048, 2017.

[59] T. Zhang and R. Baldick. Generation planning using a modified screening curve method to account for planned outage. Atlanta, GA, USA, 2019. 2019 IEEE Power & Energy Society General Meeting (PESGM), IEEE.

[60] D. Phillips, F. P. Jenkin, J. A. T. Pritchard, and K. Rybicki. A mathematical model for determining generating plant mix. Rome, Italy, 1969. Third IEEE PSCc. Rome, IEEE.

[61] K. Poncelet, E. Delarue, and W. Dh́aeseleer. Unit commitment constraints in long-term planning models: Relevance, pitfalls and the role of assumptions on flexibility. *Applied Energy, Elsevier*, 258, 2020.

[62] W. Schill. Residual load, renewable surplus generation and storage requirements in Germany. *Discussion Papers, Deutsches Institut für Wirtschaftsforschung*, 2013.

[63] Cambridge Economic Policy Associates Ltd. Study on the estimation of the value of lost load of electricity supply in Europe. Report, Agency for the cooperation of energy regulators, 2018.

[64] Y. Emre. The improved screening curve method regarding existing units. *International Journal of Electrical Power & Energy Systems, Elsevier*, 264:310–326, 2018.

[65] T. Zhang. *Generation Planning Using Screening Curve Method.* PhD thesis, University of Texas, Austin, Texas, USA, 5 2013.

[66] C. De Jonghe, E. Delarue, R. Belmans, and W. D'haeseleer. Determining optimal electricity technology mix with high level of wind power penetration. *Applied Energy, Elsevier*, 88:2231–2238, 2011.

[67] International Energy Agency (IEA) and Nuclear Energy Agency (NEA). Projecting costs of generating electricity. Report, IEA and NEA, 2015.

[68] K. Poncelet, E. Delarue, D. Six, J. Duerinck, and W. D'haeselee. Impact of the level of temporal and operational detail in energy-system planning models. *Applied Energy, Elsevier*, 162:631–643, 2016.

[69] N. E. Koltsaklis and M. C. Georgiadis. A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. *Applied Energy, Elsevier*, 158:310–331, 2015.

[70] P. Seljom, E. Rosenberg, L. E. Schaffer, and Marte Fodstad. Bidirectional linkage between a long-term energy system and a short-term power market model. *Energy, Elsevier*, 198, 2020.

[71] A. Pina, C. A. Silva, and P. Ferrao. High-resolution modeling framework for planning electricity systems with high penetration of renewables. *Applied Energy, Elsevier*, 112:215–223, 2013.

[72] F. J. de Sisternes and M. D. Webster. Optimal selection of sample weeks for approximating the net load in generation planning problems. *ESD Working Papers, MIT*, 2013.

[73] K. Poncelet, H. Hoschle, E. Delarue, A. Virag, and W. D'haeseleer. Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. *IEEE Transactions on Power Systems , IEEE*, 32:1936 – 1948, 2017.

[74] A. Almaimouni, A. Ademola-Idowu, J. N. Kutx, A. Negash, and D. Kirschen. Selecting and evaluating representative days for generation expansion planning. Dublin, Ireland, 2018. 2018 Power Systems Computation Conference (PSCC), IEEE.

[75] P. Nahmmacher, E. Schmid, L. Hirth, and B. Knopf. Carpe diem: A novel approach to select representative days for long-term power system modeling. *Energy, Elsevier*, 112:430–442, 2016.

[76] P. Vithayasrichareon, T. Lozanov, J. Riesz, and I. MacGill. Impact of operational constraints on generation portfolio planning with renewables. Denver, CO, USA, 2015. 2015 IEEE Power & Energy Society General Meeting, IEEE.

[77] A. S. Gaur, P. Das, A. Jain, R. Bhakar, and J. Mathur. Long-term energy system planning considering short-term operational constraints. *Energy Strategy Reviews, Elsevier*, 26, 2019.

[78] P. Vithayasrichareon and I. F. MacGil. Incorporating short-term operational plant constraints into assessments of future electricity generation portfolios. *Applied Energy, Elsevier*, 128:144–155, 2014.

[79] J.P. Deane, F. Gracceva, A. Chiodi, M. Gargiulo, and Brian P. Ó Gallacóir. Assessing power system security: A framework and a multi model approach. *International Journal of Electrical Power & Energy Systems, Elsevier*, 73:283–297, 2015.

[80] A. S. Brouwer, M. van den Broek, A. Seebregts, and A. Faaij. Operational flexibility and economics of power plants in future low-carbon power systems. *Applied Energy, Elsevier*, 156:107–128, 2015.

[81] J.P. Deane, A. Chiodi, M. Gargiulo, and B. P. Ó Gallacóir. Soft-linking of a power systems model to an energy systems model. *Energy, Elsevier*, 42:303–312, 2012.

[82] N. Helisto, J. Kiviluoma, and H. Holttinen. Long-term impact of variable generation and demand side flexibility on thermal power generation. *IET Renewable Power Generation*, 12:718 – 726, 2018.

[83] B. Saravanan, S. Das, S. Sikri, and D. P. Kothari. A solution to the unit commitment problem—a review. *Frontiers in Energy, Springer Link*, 7:223–236, 2020.

[84] G. E. Alvarez, M. G.Marcovecchio, and P. A. Aguirre. Security constrained unit commitment scheduling: A new MILP formulation for solving transmission constraints. *Computers & Chemical Engineering, Elsevier*, 115:455–473, 2018.

[85] DIgSILENT GmbH, Gomaringen, Germany. *DIgSILENT PowerFactory 2020 Manual*, 1 edition, 1 2020.

[86] S. Bu, F. R. Yu, and P. X. Liu. Distributed unit commitment scheduling in the future smart grid with intermittent renewable energy resources and stochastic power demands. *International Journal of Green Energy*, 2014.

[87] T. Li and M. Shahidehpour. Price-based unit commitment: a case of lagrangian relaxation versus mixed integer programming. *IEEE Transactions on Power Systems , IEEE*, 20:2015 – 2025, 2005.

[88] A. Bhardwaj, V. K. Kamboj, V. K. Shukla, and B. Sin. Unit commitment in electrical power system-a literature review. Melaka, Malaysia, 2012. 2012 IEEE International Power Engineering and Optimization Conference Melaka, Malaysia, IEEE.

[89] T. Brown, J. Horsch, and D. Schlachtberger. PyPSA: Python for Power System Analysis. *Journal of open research software, Software Sustainability Institute*, 2018.

[90] B. Palmintier and M. Webster. Impact of unit commitment constraints on generation expansion planning with renewables. Detroit, MI, USA, 2011. 2011 IEEE Power and Energy Society General Meeting, IEEE.

[91] J. Kim and K. S. Kim. Integrated model of economic generation system expansion plan for the stable operation of a power plant and the response of future electricity power demand. *Sustainability, MDPI*, 10, 2018.

[92] M. Howells, H. Rogner, N. Strachan, C. Heaps, H. Huntington, S. Kypreos, A. Hughe, S. Silveira, J. De Carolis, M. Bazillian, and A. Roehrl. OSeMOSYS, the open source energy modeling system: An introduction to its ethos, structure and development. *Energy Policy, Elsevier*, 39:5850–5870, 2011.

[93] advanved energy system analysis computer model EnergyPLAN. Markal/times. https://www.energyplan.eu/othertools/national/markaltimes/, 7 2020.

[94] J. G. Pena Balderrama, T. Alfstad, C. Taliotis, M. R. Hesamzadeh, and M. Howells. A sketch of Bolivia's potential low-carbon power system configurations. the case of applying carbon taxation and lowering financing costs. *Energies, MDPI*, 11, 2018.

[95] Y. Chung, C. Paik, and Y. J. Kim. Open source-based modeling of power plants retrofit and its application to the Korean electricity sector. *International Journal of Greenhouse Gas Control, Elsevier*, 81:21–28, 2019.

[96] V. Sridharan, O. Broad, A. Shivakumar, M. Howells, B. Boehlert, D. G. Groves, H. Rogner, C. Taliotis, J. E. Neumann, K. M. Strzepek, R. Lempert, B. Joyce, A. Huber-Lee, and R. Cervigni. Resilience of the eastern African electricity sector to climate driven changes in hydropower generation. *Nature Communications, Nature*, 2019.

[97] D. Dreier and M. Howells. OSeMOSYS-PuLP: A stochastic modeling framework for long-term energy systems modeling. *Energies, MDPI*, 12, 2019.

[98] J. Després. Modelling the long-term deployment of electricity storage in the global energy system. *Electric power. Université Grenoble Alpes*, 2015.

[99] J. Després. Development of a dispatch model of the European power system for coupling with a long-term foresight energy model. *Working Papers hal-01245554, HAL*, 2015.

[100] N. Helistö, J. Kiviluoma, H. Holttinen, J. D. Lara, and B. Hodge. Including operational aspects in the planning of power systems with large amounts of variable generation: A review of modeling approaches. *WIREs Energy and Environment, Wiley*, 8, 2019.

[101] M. van den Broek, A. Faaij, and W. Turkenburg. Planning for an electricity sector with carbon capture and storage: Case of the Netherlands. *International Journal of Greenhouse Gas Control 2, Elsevier*, pages 105–129, 2008.

[102] K. Van den Bergh, K. Bruninx, E. Delarue, and W. D́haeseleer. LUSYM: A unit commitment model formulated as a mixed-integer linear program. *TME WORKING PAPER - Energy and Environment, KU Leuven*, 2016.

[103] P. Vithayasrichareon and I. F. MacGil. A Monte-Carlo based decision-support tool for assessing generation portfolios in future carbon-constrained electricity industries. *Energy Policy, Elsevier*, 41:374–392, 2012.

[104] G. A. Duarte Pereira. Decision analysis for sustainable energy systems design: The São Miguel –Azores case study. Master's thesis, Universidade Técnica de Lisboa, Lisbon, Portugal, 10 2010.

[105] J. Kiviluoma, E. Rinne, and N. Helisto. Comparison of flexibility options to improve the value of variable power generation. *International Journal of Sustainable Energy*, 37, 2018.

[106] Q. Zhang, B. C. Mclellan, T. Tezuka, and K. N.Ishihara. An integrated model for long-term power generation planning toward future smart electricity systems. *Applied Energy, Elsevier*, 112:1424–1437, 2013.

[107] E. Delarue and A. Belderbos. Accounting for flexibility in power system planning with renewables. *International Journal of Electrical Power & Energy Systems, Elsevier*, 71:33–41, 2015.

[108] E. Delarue, D. Cattrysse, and W. D'haeseleer. Enhanced priority list unit commitment method for power systems with a high share of renewables. *Electric Power Systems Research, Elsevier*, 105:115–123, 2013.

[109] A. Mills and R. Wiser. Changes in the economic value of variable generation at high penetration levels: A pilot case study of California. Technical report, Environmental Energy Technologies Division, Ernest Orlando Lawrence Berkely National Laboratory, 2012.

[110] F. Marini and J. M. Amigo. *Hyperspectral Imaging*. Data Handling in Science and Technology, Elsevier, The Netherlands, 2020.

[111] J. H. Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 1963.

[112] M. Hoffmann, L. Kotzur, D. Stolten, and M. Robinius. A review on time series aggregation methods for energy system models. *Energies, MDPI*, 13, 2020.

[113] K. Van den Bergh and E. Delarue. Cycling of conventional power plants: Technical limits and actual costs. *Energy Conversion and Management, Elsevier*, 97:70–77, 2015.

[114] Frankfurt Institute of Advanced Studies Renewable Energy Group. powerplantmatching toolset. https://github.com/FRESNA/powerplantmatching, 7 2020.

[115] D. Gearino. Inside clean energy: A California utility announces 770 megawatts of battery storage. That's a lot. https://insideclimatenews.org/news/06052020/inside-clean-energy-california-battery-storage-wind-solar, 7 2020.

[116] R. Chourishi. Image from the cover page. https://www.electricalindia.in/artificial-intelligence-an-advanced-approach-in-power-systems/, 7 2020.

[117] European Network of Transmission System Operators for Electricity. Transparency Platform. https://transparency.entsoe.eu/, 7 2020.

[118] M. Joos and I. Staffell. Short-term integration costs of variable renewable energy: Wind curtailment and balancing in Britain and Germany. *Renewable and Sustainable Energy Reviews, Elsevier*, 86:45–65, 2018.

A
Appendix

A.1. Simulation preparation in OSeMOSYS

This section of the Appendix is complementary to Case study I (see Section 4.3), Case study II (see Section 4.4) and Case study III (see Section 4.5) in which OSeMOSYS is used as a static model for generation expansion planning (see Subsection 2.2.3). In order to do so, there are a few concepts and changes that have to be introduced beforehand. These ones are herein specified:

A.1.1. Approximated LDC in OSeMOSYS

OSeMOSYS offers two parameters to insert timeslices and their corresponding weights: **SpecifiedDemand-Profile** and **YearSplit**, respectively. For timeslice u, $YearSplit_u = \omega_u$. This weight can be directly mapped to the timeslice duration, i.e. $d_u = \omega_u \cdot 8760$. On the other hand, the **SpecifiedDemandProfile** refers to fractions of the total electrical energy E_T to be supplied. The timeslices obtained by the Ward agglomerative hierarchical clustering algorithm have units of power (e.g. MW) so they must be multiplied by their corresponding duration d_u to obtain their value e_u in units of energy. Then, each timeslice expressed in energy units is divided by E_T to obtain its corresponding energy fraction, $SpecifiedDemandProfile_u = e_u/E_T$.

A.1.2. Standard screen curve in OSeMOSYS

In order to recreate the results obtained with the standard screen curve (SSC) method (Subsection 2.2.3) in OSeMOSYS, the code needs to be slightly modified. When only one year is considered in the program, most variables are automatically not discounted. Nonetheless, the **DiscountedOperatingCost** is in fact discounted and its value for a given year ($current\_year$) is computed in reference to an initial year ($initial\_year$) as follows:

$$DiscountedOperatingCost = OperatingCost \cdot \frac{1}{(1 + DiscountRate)^{(current\_year - initial\_year + opc\_value)}} \tag{A.1}$$

```
model += DiscountedOperatingCost[r][t][y] == OperatingCost[r][t][y] * (1 /((1 +
    DiscountRate[r]) ** (int(y) - min([int(yy) for yy in YEAR]) + opc_value)))
```

Listing A.1: Excerpt from OSeMOSYS-PuLp that shows the computation of the **DiscountedOperatingCost**. The **OperationCost** is not discounted when $opc\_value = 0$.

The opc_value is originally $opc\_value = 1/2$; therefore, when $current\_year = initial\_year$ (which is the case when only a single year is considered), the **OperatingCost** gets discounted. Since the SSC method does not consider this discount, the opc_value needs to be set to $opc\_value = 0$ to obtain the same results.

The second adjustment needed is to eliminate the **SalvageValue**. To do this artificially, the first step is to set the **DepreciationMethod** equal to 2. In OSeMOSYS, this means that the assets' depreciation is calculated based on a straight-line depreciation method (that is, an asset value falls in a straight line over its lifetime). When such depreciation is selected, the **SalvageValue** in a given year ($current\_year$) is computed in reference to the final year ($final\_year$) as follows:

$$SalvageValue = CapitalCost \cdot NewCapacity \cdot \left(1 - \frac{final\_year - current\_year + 1}{OperationalLife}\right) \tag{A.2}$$

```
model += SalvageValue[r][t][y] == CapitalCost[r][t][y] * NewCapacity[r][t][y] * (1 -
     (max([int(yy) for yy in YEAR]) - int(y) + 1) / OperationalLife[r][t])
```

Listing A.2: Excerpt from OSeMOSYS-PuLp that shows the computation of the **SalvageValue**. The **SalvageValue** is artificially set to $SalvageValue = 0$ when $OperationalLife = 1$.

When $final\_year = current\_year$, the only way to eliminate the **SalvageValue** is to set the operational lifetime of each technology to $OperationalLife = 1$. In summary, the needed modifications to recreate the standard screen curve (SSC) method in OSeMOSYS are: $opc\_value = 1/2$, $DepreciationMethod = 2$, $OperationalLife = 1$ (for all technologies).

A.2. PowerFactory interface

This section of the Appendix is complementary to the description of how to access PowerFactory while in bidirectional softlinking with OSeMOSYS and PyPSA (see Subsection 3.3.2).

```

def char_load(element, path_to_file, file_name, parameter , year_counter):

    #Access the element in PF
    elmt=app.GetCalcRelevantObjects(element.ident)[0]

    #Look for all existing characteristics and delete them
    chaOld = elmt.GetContents('*.Cha*')

    for cha in chaOld:
        cha.Delete()
```

```
    #access the subfolder 'Characteristics' in the 'Operational Library'
    charFolder=app.GetProjectFolder('chars')

    #create an object of the type 'ChaTime'
    newChar=charFolder.CreateObject('ChaTime', element.name+'Char')

    #define the parameters of the new characteristics
    newChar.source=1 #external file to be the source
    newChar.iopt_stamp=0 #the external file is not time stamped
    newChar.usage=2 #values are absolute values
    newChar.fileunit=1 #the file units are in 1h
    newChar.repeat=3 # recurrance of the data
    newChar.datacol=year_counter+2
    #this assumes the year_counter starts at 0
    #time is in column 1, so the year_counter is shifted by 2

    #year_counter is for characteristics that are created at once (e.g. load profile for 50 years),
    #but need to be accessed iteratively in a loop

    #Path to csv file
    newChar.f_name=path_to_file+"/"+file_name
    newChar.ioptfile=0 #it is a csv file

    #Create ChaRef object and name it parameter
    #(e.g. plini, the active power consumption of the load)
    refObj=elmt.CreateObject('ChaRef', parameter)
    #Assign created ChaTime to ChaRef
    refObj.typ_id=newChar
```

Listing A.3: Example of a **Loader**. This piece of code loads a specific characteristic (e.g. timeseries) to a component (e.g. load object) of the network.

```

def result_extract(pf_monitored_variables, gen_type):

    results=app.GetFromStudyCase(cg.pf_results_folder)
    # results folder from DC UC in PF: "Unit Commitment (after optimisation) DC.ElmRes"
    results.Load()
    num_row=results.GetNumberOfRows()
    num_col=sum(map(len, pf_monitored_variables.values()))
    # count the number of variables in the dictionary

    result_matrix= np.zeros([num_row, num_col])

    i=-1

    for gen in pf_monitored_variables:
        gen_str=gen+gen_type
```

```
        gen_temp=app.GetCalcRelevantObjects(gen_str)

        for var in pf_monitored_variables[gen]:
            col_index=results.FindColumn(gen_temp[0],var)
            i=i+1

            for row in range(num_row):
                value=results.GetValue(row, col_index)
                # get value one row at a time (only possibility in PF)
                if abs(value[1])<1*10**-5: # eliminate numerical error
                    result_matrix[row, i]=0
                else:
                    result_matrix[row, i]=value[1]

    return result_matrix

    # Columns in result_matrix follow the same order as in the input (ordered) dictionary. For instance,
    # if the first key has 5 values, the first 5 columns of result_matrix will be for the values of that key.
```

Listing A.4: Example of an **Extractor**. This piece of code extracts the results from all variables specified in the dictionary *pf_monitored_variables*.

A.3. Details about Case study III

This section of the Appendix is complementary to Case III, where a multi-generator system with renewable energy investment and storage arbitrage is studied (see Section 4.5). The context for this system is the Dutch electricity sector. The details are herein specified:

1. System and general:
    - The investment year considered is 2030.
    - The electric demand corresponds to the Dutch load profile with a power peak of 18.94 GW, as per 2018 [117].
    - The transmission system is a closed system and a copper plate.
    - The cost to curtail energy from renewable sources is assumed to be 92.32 EUR/MWh, which corresponds to the cost paid out due to wind curtailment in Germany in 2016 [118].
2. Conventional generation:
    - The conventional power plants existing in the Netherlands are obtained from Ref. [114]. This results in a total of 50 power plants.
    - The technology types (i.e. nuclear, coal and CCGT) are specified for the vast majority of power plants in Ref. [114]. However, the technology COCGT is assigned to those ones without a specific technology type.
    - The installed capacity of each power plant is re-expressed as a ratio to the total installed capacity of their corresponding technology group.

- After OSeMOSYS optimizes the energy mix based on technologies, the resulting installed capacity per technology is assigned to their corresponding group of power plants. Thus, each power plant obtains an installed capacity that is proportional to its aforementioned ratio.
- As there are no specific plans in the Netherlands to increase the share of nuclear in the mix (e.g. in Ref. [8]), this technology is not considered as an investment option in OSeMOSYS. However, the existing 485 MW Borselle nuclear power plant is still taken into consideration.
- As per projected political decisions given by the Dutch Climate Agreement [8], coal shall no longer be a part of the energy mix by 2030. Hence, all existing coal power plants are kept, but their installed capacities are set to zero. In OSeMOSYS, coal is no longer considered as an investment option.
- In Ref. [114], the marginal cost of each power plant is not specified. Hence, it is assumed that these costs are Gaussian-distributed for a certain technology type. The mean of the Gaussian curve is given by the paremeters defined in Table 2.2, while the standard deviation is set to $\sigma = 2.5$.
- All these generators form the electricity price, according to an energy-only market (see Subsection 2.2.3).
- The unit commitment (UC) problem is ignored. This means that binary variables are disregarded and only the economic dispatch (ED) is run. UC becomes too computational demanding for 50 power plants.

3. Renewable energy sources:
   - The renewable energy (R.E.) sources considered are solar, onshore wind and offshore wind. Their 2018 profiles are obtained from Ref. [117].
   - These profiles are normalized (i.e. divided by their maximum value). Subsequently, they are concatenated with the normalized load profile, forming a single matrix. This matrix is passed through the process specified in Section 4.2 to select the load timeslices.
   - The load timeslices are in the same row as the R.E. timeslices. These R.E. timeslices are in fact the capacity factors of the R.E. sources at the selected load timeslices.
   - These capacity factors are passed to OSeMOSYS parameter: **CapacityFactor**.
4. Storage units:
   - The number of storage units are pre-selected and are thus not considered as investment options. The storage units selected are batteries. The number of batteries is set to 10.
   - The installed capacities i_{bat}^{cap} of these batteries are randomly chosen integers. The specific vector is: $i_{bat}^{cap} = [50, 65, 70, 75, 100, 110, 115, 125, 130, 150]$ MW.
   - Their energy capacity e_{bat}^{cap} is equivalent to four hours of operation at full power, i.e. $e_{bat}^{cap} = [200, 260, 280, 300, 400, 440, 460, 500, 520, 600]$ MWh.
   - These power and energy sizes are justified as per actual, planned projects. Recently, Southern California Edison announced seven battery projects with power capacity within the range $i_{bat}^{cap} = [50, 230]$ MW. The McCoy project with $i_{bat}^{cap} = 230$ MW will have an energy capacity of $e_{bat}^{cap} = 920$ MWh. The seven projects are expected to be completed by August, 2021 [115].
   - The marginal cost of each of these batteries is set to the electricity price. In this way, batteries store energy when the price is low and inject energy when the price is high (i.e. arbitrage).