THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

COMPOSANTS ET SYSTEMES DE GESTION DE L'ENERGIE

Présentée et soutenue par :

M. HUGO RADET

le jeudi 3 mars 2022

Titre :

Integrated design methods for distributed multi-energy systems under uncertainties

---

Ecole doctorale :

Génie Electrique, Electronique, Télécommunications (GEETS)

Unité de recherche :

Laboratoire Plasma et Conversion d'Energie (LAPLACE)

Directeur(s) de Thèse :

M. XAVIER ROBOAM

M. BRUNO SARENI

Rapporteurs :

MME FLORENCE OSSART, UNIVERSITE SORBONNE

M. ROBIN ROCHE, UNIV TECHNO BELFORT MONTBELIAR

Membre(s) du jury :

M. BRUNO FRANCOIS, ECOLE CENTRALE DE LILLE, Président

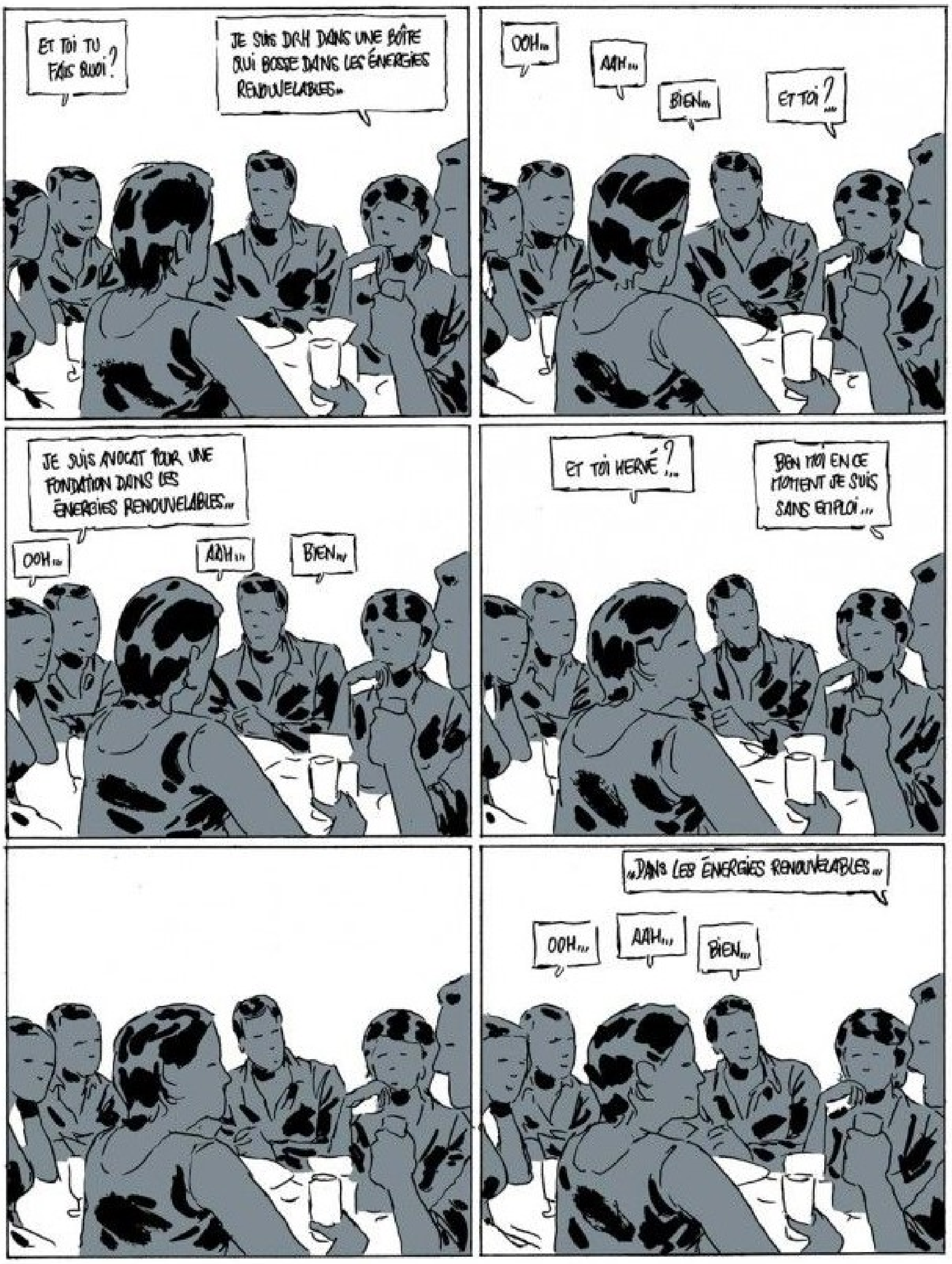
M. BRUNO SARENI, TOULOUSE INP, Membre

M. MAURIZIO REPETTO, POLITECNICO DE TURIN, Membre

M. PIERRE HAESSIG, SUPELEC RENNES, Membre

M. VINCENT DEBUSSCHERE, INP GRENOBLE, Membre

M. XAVIER ROBOAM, TOULOUSE INP, Membre



Fabcaro, *Open bar 2*

Acknowledgment

A l'heure où je rédige ces remerciements, quelques semaines se sont écoulées depuis la soutenance. L'impatience d'en finir a laissé place à une pointe de nostalgie. Etrange sentiment que la fin !

Mes premiers remerciements s'adressent naturellement à mes directeurs de thèse Bruno Sareni et Xavier Roboam, sans qui rien n'aurait été possible. Merci pour le temps et l'énergie (renouvelable !) que vous avez consacrés à cette thèse. Je vous remercie également pour la confiance et l'autonomie que vous m'avez accordées dans ces aventures mathématico-conceptuelles dont je maitrisais peu l'issue…

Ce premier contact avec Toulouse n'aurait pas vu le jour sans Rémy Rigo-Mariani. Merci à toi pour tes conseils précieux, ton introduction aux arcanes du monde des "MILP-eux" et ton travail inspirant. Merci également à Pierre Haessig, qui en plus d'avoir accepté de faire partie de mon jury, a stimulé ma curiosité et confirmé mon engagement dans ces travaux.

Un grand merci à l'ensemble des membres du jury pour leurs questions "un peu sioux" et la richesse des discussions qui ont suivies : Florence Ossart, Robin Roche, Bruno François, Pierre Haessig, Maurizio Repetto et Vincent Debusschere. Merci à vous d'avoir fait le déplacement !

Ces trois années auront été perturbées par une invitée surprise : la Covid. De petite taille mais terriblement impactante sur la vie de tous les jours. Malgré le peu de temps passé au labo, je souhaitais également remercier l'ensemble des membres de l'équipe Genesys. En particulier, Eric Bru, Nicolas Roux, Henri Schneider et Fabien Lacressonnière avec qui j'ai eu plaisir à travailler.

Comment ne pas avoir un mot pour l'ensemble des professeurs qui ont éveillé ma curiosité scientifique – et pas seulement - tout au long de ma scolarité. Une mention spéciale pour Philippe Lieutaud et Roch Cassanas qui ont marqué plusieurs générations d'étudiants dont je fais partie. Mille mercis !

Et puis les amis… Ô combien nécessaires à l'édifice ! Je pense en particulier à ceux du début : Romain, Mickael, Yannis et Alexis. Ceux du lycée : Rémi, Thibaut, Fanny, Pauline. Les grelous : Arthur, Mathieu, Pierre, Salomé, Marion, Grégoire, Maud, Max, Istas. Et les inclassables : Vincent, Tonio, Ben, Louis, Emile, Fab, Jean et Manu. Merci pour tous ces bons moments partagés à grimper, discuter, rigoler, picoler, turlupiner, voyager, (et bientôt...) pêcher !

Enfin, les mots ne suffisent pas pour exprimer mon immense gratitude envers ma famille, mes parents et ma sœur pour tout ce qu'ils m'ont apporté. Merci pour votre amour, votre bienveillance, et votre soutien sans faille. Sans vous, rien de tout cela n'aurait vu le jour... Merci infiniment !

En toute originalité, mes derniers mots s'adressent à Olivia. Une sorte d'hybride entre Aphrodite et Dionysos... Merci pour l'ivresse amoureuse du quotidien !

Hugo - Avril 2022

Abstract

With the growing integration of variable renewable energy (VRE) sources into the conventional power grid, the concept of distributed energy systems (DES) has emerged: the energy is produced close to the end-user, and flexibility means are included in the system to supply energy demands at any time. Among the considered options, storage systems along with multi-energy strategies tend to be promising directions to mitigate the production variability by coupling the energy carriers with each other.

Planning the design of such systems is a challenging task because the problem displays multiple facets that are difficult for policy- and decision-makers to address in a systemic manner. Also, decisions are made while many parameters remain uncertain (e.g., future investment costs, energy prices, demands and production) as their values progressively unfold over time. Therefore, mathematical tools are often needed to provide decision support regarding several techno-economic requirements: the problem is usually expressed in the form of an optimization problem where decision variables are the sizes of the equipment.

This work addresses this issue by developing a generic framework to assess and compare different design and operation strategies for multi-energy systems. Then, three critical questions are tackled using this framework. In the first part of the thesis, the deterministic design model is built. Solving such a model is fast and allows running parametric analysis to assess the value of multi-energy systems and seasonal storage to supply residential customers with a high share of solar production.

Then, the second part of this work addresses the design of DES under uncertainty. To this end, two design methods based on stochastic programming are developed: one relies on mathematical programming and the other uses a metaheuristic algorithm. To solve the problem in a reasonable time, these methods are usually based on simplified versions of the problem. In particular, sizing values are computed assuming perfect foresight of the operation strategy for a given scenario. The main objective of this part is to challenge this hypothesis by jointly evaluating the design solutions with realistic operation policies which only have access to past and current information. In addition, this work aims at further investigating the close relationship between operation and design. Should the operation strategy used to design the system and the one used in real-time be strictly identical? This part attempts to clarify this point.

Finally, the last part of this work deals with the dynamic design of DES. In this case, the model takes technology replacement due to aging into account, so multiple design decisions must be made over the horizon. Unlike the majority of studies, the optimization model includes the impact of the operation over system lifetimes: the latter are not fixed *a priori*, but they depend on the way technologies are operated over time. The aware aging method (which comes from the literature) is then compared with two heuristic design strategies based on single representative years.

All the previous methodological developments are applied to a DES which may include a set of hydrogen units (i.e., fuel cell, electrolyzer and storage tank) where the cogenerated heat can be recovered to supply thermal energy demands.

Résumé

Ces dernières années, le développement accéléré des énergies renouvelables a permis d'envisager de nouveaux modèles énergétiques basés sur des modes de production décentralisés : l'énergie est produite proche du consommateur et des moyens de flexibilité sont installés pour garantir, à tout instant, l'équilibre entre la consommation et la production. Parmi l'ensemble des solutions envisagées, le stockage est un moyen privilégié pour pallier l'intermittence de la production. Ce dernier peut s'accompagner de stratégie "multi-énergies" qui permettent d'améliorer la performance globale du système en couplant plusieurs vecteurs énergétiques entre eux.

La conception de tels systèmes est un problème complexe car il nécessite l'appréhension d'un grand nombre de paramètres, et ce de manière systémique. Qui plus est, les décisions de dimensionnement doivent être prises alors que de nombreux paramètres sont incertains (e.g., coûts d'investissement, prix de l'énergie, consommation et production futures). Par conséquent, le problème de conception est généralement formulé sous la forme d'un problème d'optimisation où les dimensions des équipements sont les variables de décision du problème à résoudre.

La thèse s'inscrit dans ce contexte en développant des stratégies de conception et de pilotage pour des systèmes multi-énergies. En particulier, le travail s'articule autour de trois axes de recherche principaux. La première partie de la thèse a pour objectif de développer un modèle d'optimisation du dimensionnement dans un cadre déterministe. La résolution d'un tel problème est rapide, ce qui permet d'effectuer des analyses paramétriques. De cette façon, cette partie interroge la pertinence technico-économique de tels systèmes pour alimenter des consommateurs résidentiels avec une part d'énergie renouvelable croissante.

Ensuite, la deuxième partie de ce travail (qui constitue le cœur de la thèse) s'intéresse au problème de conception sous incertitudes. Pour ce faire, deux méthodes de dimensionnement basées sur la programmation stochastique sont introduites : l'une est basée sur la programmation linéaire et l'autre utilise une métaheuristique. De façon à résoudre le problème d'optimisation en un temps imparti, ces méthodes sont généralement basées sur des versions simplifiées du problème. En particulier, les dimensions des systèmes sont calculées en considérant que la stratégie de pilotage a une connaissance parfaite du futur. L'objectif de cette partie est donc de questionner cette hypothèse en évaluant conjointement les résultats de dimensionnement avec des méthodes de pilotage réalistes qui

n'anticipent pas le futur. Ce travail permet également d'étudier plus précisément le lien étroit entre la performance de la loi de gestion et le dimensionnement. Faut-il obligatoirement utiliser la loi de gestion réelle en phase de dimensionnement ? Cette partie tente d'apporter quelques réponses à cette question.

Enfin, la dernière partie de ce travail s'intéresse au problème de dimensionnement dynamique : les systèmes vieillissent et plusieurs décisions de dimensionnement doivent être prises sur l'horizon de l'étude. Contrairement à la plupart des études de la littérature, ce travail introduit l'impact de la loi de gestion sur la durée de vie des systèmes : cette dernière n'est pas fixée *a priori*, mais dépend de la façon dont le système est piloté au cours du temps. Cette méthode de dimensionnement originale intégrant le vieillissement (qui provient de la littérature) est ensuite comparée à deux heuristiques de conception basées sur une année équivalente.

Pour illustrer l'ensemble de ces travaux, les méthodes sont appliquées à un système décentralisé qui peut inclure des technologies hydrogène (i.e., pile à combustible, électrolyseur et stockage) où la chaleur produite par cogénération est valorisée pour alimenter des besoins en chaud.

Contents

List of publications

The content of chapter 3 will be presented in the Electrimacs 2022 conference (in May 2022):

- **Radet, H.**, Sareni, B., Roboam, X., 2022. Generation of energy demand and PV production profiles based on Markov chains for the design and operation of micro-grids. 14th International Conference of the International Association for Mathematics and Computer in Simulation.

Chapter 4 is based on the following (under-review) journal papers:

- **Radet, H.**, Roboam, X., Sareni, B., Design under uncertainties of a multi-energy system with seasonal storage: on the importance of the operation strategy in the design procedure (Submitted to Energy)

- **Radet, H.**, Sareni, B., Roboam, X., On the interaction between the design and operation under uncertainties of a simple distributed energy system (Submitted to COMPEL)

The third paper was also presented at the Optimization and Inverse Problems in Electromagnetism (OIPE) conference:

- **Radet, H.**, Roboam, X., Sareni, B, 2021. Method for the design and control under uncertainties of a simple distributed energy system. 16th International Workshop on Optimization and Inverse Problems in Electromagnetism (OIPE)

Chapter 5 is based on the following journal paper, which was presented at the Symposium du Génie Electrique (SGE) conference:

- **Radet, H.**, Roboam, X., Sareni, B, Rigo-Mariani, R., 2021. Dynamic ware aging design of a simple distributed energy system: a comparative approach with single stage design strategies. Renewable and Sustainable Energy Reviews 147, 111104. https://doi.org/10.1016/j.rser.2021.111104

- **Radet, H.**, Roboam, X., Sareni, B, Rigo-Mariani, R., 2021. A multi-stage design framework for the investment and operation of a simple microgrid: a comprehensive approach. Symposium du Génie Electrique (SGE)

At the beginning of my PhD, I also co-supervised a student who presented her work at the European Conference on Power Electronics and Application (EPE'2020 ECCE Europe):

- Birou, C., Roboam, X., **Radet, H.**, Lacressonnière, F., 2020. Techno-economic analysis of second-life lithium-ion batteries integration in microgrids. 22nd European Conference on Power Electronics and Applications (EPE'20 ECCE Europe) https://doi.org/10.23919/EPE20ECCEEurope43536.2020.92156

Notations

Acronyms

Notation	Description
CVaR	Conditional Value At Risk
DES	Distributed Energy System
DMES	Distributed Multi-Energy System
EAC	Equivalent Annual Cost
GA	Genetic Algorithm
LP	Linear Programming
MILP	Mixed Integer Linear Programming
MPC	Model Predictive Control
NPV	Net Present Value
OLFC	Open-Loop Feedback Control
PEME	Proton-Exchange Membrane Electrolyzer
PEMFC	Proton-Exchange Membrane Fuel Cell
PSO	Particle Swarm Optimization
PV	Photovoltaic
RB	Rule-Based
SoC	State-Of-Charge
SoH	State-Of-Health
TES	Thermal Energy Storage
VRE	Variable Renewable Energy

Table 1: Acronyms

Sets

Notation	Description
$h \in \{1, \ldots, H\}$	Set of hours
$y \in \{1, \ldots, Y\}$	Set of years
$s \in \{1, \ldots, S\}$	Set of scenarios

Table 2: Model sets

Design variables

Notation	Description	Unit
$E_y^{b,d}$	Li-ion maximum capacity in year y	[kWh]
$E_y^{tes,d}$	TES maximum capacity in year y	[kWh]
$E_y^{tk,d}$	H2 tank maximum capacity in year y	[kWh]
$p_y^{fc,d}$	PEMFC maximum power in year y	[kW]
$p_y^{el,d}$	PEME maximum power in year y	[kW]
$p_y^{pv,d}$	PV peak power in year y	[kWp]
$E_y^{b,state}$	Li-ion state design variable in year y	[kWh]
$p_y^{pv,state}$	PV state design variable in year y	[kWp]

Table 3: Model design variables

Operation variables

Notation	Description	Unit
$p_{h,y}^{b,+}$, $p_{h,y}^{b,-}$	Li-ion charge/discharge power in hour h and year y	[kW]
$p_{h,y}^{tes,+}$, $p_{h,y}^{tes,-}$	TES charge/discharge power in hour h and year y	[kW]
$p_{h,y}^{tk,+}$, $p_{h,y}^{tk,-}$	H2 tank charge/discharge power in hour h and year y	[kW]
$p_{h,y}^{fc,e}$, $p_{h,y}^{fc,h,y}$, $p_{h,y}^{fc,h2}$	PEMFC elec., heating and H2 power in hour h and year y	[kW]
$p_{h,y}^{el,e}$, $p_{h,y}^{el,h}$, $p_{h,y}^{el,h2}$	PEME elec., heating and H2 power in hour h and year y	[kW]
$p_{h,y}^{ht,e}$, $p_{h,y}^{ht,h}$	Heater electrical and heating power in hour h and year y	[kW]
$p_{h,y}^{g,+}$, $p_{h,y}^{g,-}$	Power from/to the grid in hour h and year y	[kW]
$E_{h,y}^{b}$	Li-ion state of charge in hour h and year y	[kWh]
$E_{h,y}^{tes}$	TES state of charge in hour h and year y	[kWh]
$E_{h,y}^{tk}$	H2 tank state of charge in hour h and year y	[kWh]
$A_{h,y}^{b}$	Li-ion state of health in hour h and year y	[kWh]

Table 4: Model operation variables

Parameters

Notation	Description	Unit
r	Discount rate	[0,1]
γ	Annuity factor for each technology	[0,1]
Δh	Hourly time step	[h]
Δy	Yearly time step	[y]
η^-, η^+	Storage charge/discharge efficiencies	[0,1]
η^{loss}	Storage self-discharge coefficient	[h^{-1}]
$\underline{e}$, $\overline{e}$	Storage lower/upper state bound factors	[0,1]
$\underline{p}$, $\overline{p}$	Storage lower/upper power bound factors	[h^{-1}]
$\eta^{ht,e\rightarrow h}$	Heater conversion efficiency	[0,1]
$\eta^{fc,h_2\rightarrow e}$, $\eta^{fc,h_2\rightarrow h}$	PEMFC conversion efficiencies	[0,1]
$\eta^{el,e\rightarrow h2}$, $\eta^{el,e\rightarrow h}$	PEME conversion efficiencies	[0,1]
$\overline{g}$	Maximum grid power	[kW]
$p^{ld,e}_{h,y}$, $p^{ld,h}_{h,y}$	Elec. and thermal demand in hour h and year y	[kW]
$p^{pv}_{h,y}$	PV capacity factor in hour h and year y	[0,1]
τ^{sh}	Share of solar production	[0,1]
c^b_y	Li-ion investment cost in year y	[€/kWh]
c^{tes}_y	TES investment cost in year y	[€/kWh]
c^{tk}_y	H2 tank investment cost in year y	[€/kWh]
c^{fc}_y	PEMFC investment cost in year y	[€/kW]
c^{el}_y	PEME investment cost in year y	[€/kW]
c^{pv}_y	PV investment cost in year y	[€/kWp]
$c^{g,+}_{h,y}$, $c^{g,-}_{h,y}$	Electricity rate and feed-in tariff in hour h and year y	[€/kWh]
π_s	Probability of scenario s	[0,1]

Table 5: Model parameters

Chapter 1

Introduction

Highlights

- This thesis revolves around four key points:
  1. What is the additional value brought by multi-energy systems and seasonal storage to supply customers with a high share of solar production? In particular, what is the contribution of the hydrogen cogenerated heat?
  2. How can uncertainties be included in energy planning studies? What kind of information is available (at each time step) to make both operation and design decisions? How does it impact the results?
  3. What about the interaction between the design and operation strategies? Should the operation strategy used to design the system and the one used in real-time be identical?
  4. How can the impact of the operation on equipment aging be included in dynamic planning models? Does it bring more value compared to design strategies based on ”snapshot” investment?
- These multifaceted issues are complex to address in their entirety in only 3 years. Thus, this work mainly focuses on methodological aspects attached to design strategies.
- Technological details are kept deliberately simple as the objective is to illustrate design and operation issues rather than giving extensive quantitative results.

Contents

1.1 Context

1.1.1 Time for ”deep decarbonization”

The conclusion of the working group I contribution to the sixth assessment report from the International Panel on Climate Change (IPCC) is crystal clear: ”limiting human-induced global warming to a specific level requires limiting cumulative CO2 emissions, **reaching at least net zero CO2 emissions**, along with strong reductions in other greenhouse gas (GHG) emissions” [1]. The message is far from new but it becomes more than urgent today to avoid the worst impacts of climate change. As a result, a growing number of national governments have announced net-zero emission pledges by 2050 but a few have made this a legal obligation [2]. Among the different levers for reducing anthropogenic GHG emission, there is a widespread agreement that electricity will play a major role along with the massive development of variable renewable energies (VRE) [3, 4]. However, the variable and distributed nature of supply (i.e., unlike centralized generators, VRE units are usually distributed throughout the country) poses new challenges to the power system. On one hand, new flexibility sources are needed to synchronize production and demand. On the other hand, the power grid has not been designed to cope with the high fluctuation and increasing complexity of flows due to the emergence of many new producers.

1.1.2 The role of distributed energy systems

Within this context, the concept of distributed energy systems [1] (DES) has emerged: the energy is produced close to the end-user and the system includes several technologies that convert, store and deliver energy to supply the demands at any time [5]. To increase its

[1] Microgrids are part of the DES category. Their specificity is that they can be disconnected from the main power grid based on techno-economic requirements.

flexibility, a DES may have the ability to manage its consumption [6]. It can also include multiple energy carriers with the technologies needed to exchange energy between nodes [7]. Coupling the energy carriers with each other seems to be a promising direction to increase the system performance and mitigate production variability [8].

As a result, DES can be valuable assets for grid operators as they can provide the flexibility required by the traditional system. However, DES are mostly operated by private entities and cannot be directly controlled by grid operators. Thus, new market designs are emerging to encourage DES owners to participate in the grid regulation effort. In this way, the interests of both grid operators and DES owners are satisfied: the former ensure a smooth operation of the grid while the latter are rewarded for the value they brought [9].

1.1.3 The Hymazonie project

This work is part of the Hymazonie project supported by the French Agency for Ecological Transition (ADEME). The partners include industrial and academic players such as CNES (French National Centre for Space Studies), CEA (French Alternative Energies and Atomic Energy Commission), LAPLACE (Laboratoire Plasma et Conversion d'Energie), GDI (Guyane Developpement Innovation), and the University of French Guiana. The project aims at studying the feasibility of a multi-energy system (MES) including hydrogen units that co-generate both electricity and heat to supply energy demands. To this end, a prototype will be installed on the university's Troubiran campus in Cayenne where various applications will be tested. At the same time, another objective of the project has been to **develop a modeling framework to design and assess the techno-economic performance of such a system**. This was the goal of the LAPLACE laboratory and the reason of this work.

Initially, these two facets should have been run simultaneously, allowing back and forth between the simulation and real-life experiments. Unfortunately, the project did not start as expected and this thesis has been done without any connection to the Hymazonie case study. Therefore, it was decided to broaden the subject to the general case of distributed multi-energy systems (DMES) (the Hymazonie prototype is in a way a DMES) that can be connected to external networks (see figure 1.1 for a schematic representation). In the following, the DMES is assumed to be operated in a centralized manner and distributed control strategies will not be addressed in this work.

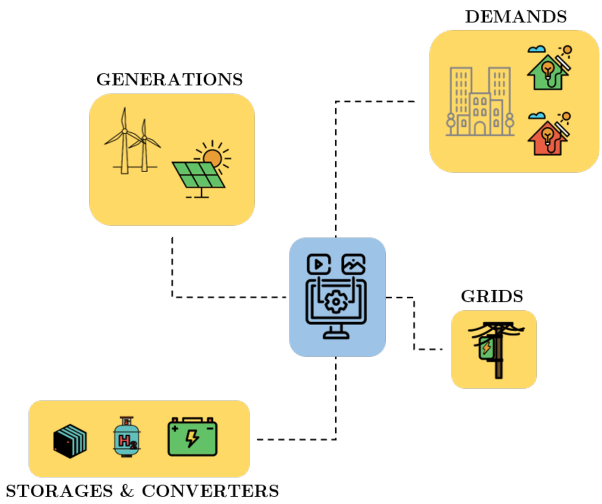


Figure 1.1: Schematic view of a DMES. The system is assumed to be operated in a centralized manner.

1.1.4 The Ausgrid dataset

As no data were available, case studies were tailored from the literature to illustrate the methodological points developed in this work. The methods are generic in the sense that the same approaches can easily be used once the Hymazonie parameters become available.

Therefore, the following case studies are based on the Ausgrid (Australian distributor of electricity) dataset [10] where 3 years of measured energy demands and production time series (at a 30 min time step) are openly available for 300 residential customers. The authors in [11] extracted a ”clean dataset” from which 20 customers are identified with clean solar production profiles along with both electrical and thermal demand. Note that the thermal energy consumption corresponds to water heating without any information about the final usage (domestic hot water? space heating? both?). From this reduced dataset, 60 synchronized hourly profiles over 1 year (20 customers over 3 years where measurements are resampled at an hourly time step) are then used to illustrate the methodological points developed in this work [2]. Figure 1.2 shows one week of energy demands and production for the 39th customer as an example. Each part of this work is associated with a case study derived from this dataset but tailored according to the methodological aspect that has to be highlighted.

[2]In the rest of the study, the 60 profiles only represent residential customers regardless of where they come from. In general in this work, customers profiles are used along with the french tariff of electricity.

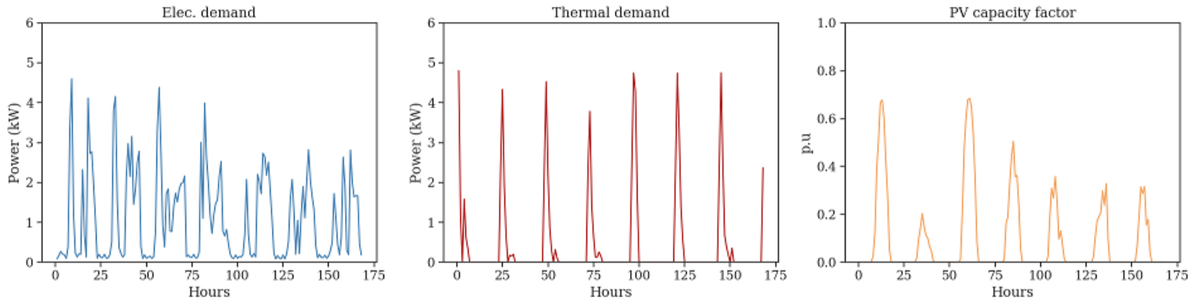


Figure 1.2: Example of one week of energy demands and production for the 39th customer of the Ausgrid dataset.

1.2 Motivations

Planning the design of such distributed energy systems is a challenging task because the problem displays multiple facets that are difficult for policy- and decision-makers to address in a systemic manner. In particular, given long-term horizons, decisions have to be made while many parameters remain uncertain as their values progressively unfold over time. Therefore, decision-making tools are often needed to provide decision support regarding several techno-economic requirements. The problem is most of the time expressed in the form of an optimization problem where decision variables are the equipment sizes and the power flows controlled in the DES. A large number of energy modeling tools are already available and many efforts have been made in the literature to review the different technical and methodological aspects of each approach [12, 13, 14, 15]. The objective of this section is not to repeat what can be found in the aforementioned references, but rather give a general overview of the issue to help the reader understand the scope of this work. More detailed states of the arts are given at the beginning of each chapter for each research question addressed in this thesis.

1.2.1 Main ingredients

The objective of this section is to introduce the main ingredients of an energy modeling problem as they will form the basis of this work.

Traditionally, the literature starts by introducing the differences between optimization models used to compute the size of the assets and their simulation counterparts [13, 14, 16, 17]. This classification gives the impression that they are two distinct worlds with little overlap. This section attempts to present the problem differently as it seems to the author more appropriate to understand the rest of this work. The objective is to get a broader perspective on the energy modeling issue by reconnecting, in a certain way, the two worlds (i.e., optimization and simulation). Therefore, following this logic, the main ingredients of an energy modeling problem are:

- A **design strategy** to select and properly size the assets at each design time step.
- An **operation strategy** to compute the power flows at each operation time step.
- A **simulator** to evaluate both the design and operation strategies and compute the techno-economic indicators.

For the sake of generality, a clear difference is made between the model used to design the system and the one used in the simulator to compute the techno-economic indicators of a design solution. The reason for this is to allow for different granularity between the simulation and design models. As further argued in the next paragraphs, both the operation and design strategies might be based on mathematical optimization as their objective is to find "good" decisions, according to a set of constraints and objectives [3]. However, the resolution of such optimization problems often requires inevitable simplifications. Dividing the problem in this way allows for a strict distinction between the simulator, which is a good representation of the real problem, and the design and operation strategies that might be built on simplified versions of the problem. The techno-economic indicators are computed afterward by evaluating the design and operation strategies on the simulator [4]. This way, simulating both the design and operation strategies allows validating that the simplifications - made for the purpose of optimization - give consistent decisions with the requirements of a given project. In other words, **simulation** models are used to **evaluate** solutions, while **optimization** models are used to **find** them. Both models are included in the modeling framework of this thesis to avoid any loss of generality. The latter is depicted in figure 1.3.

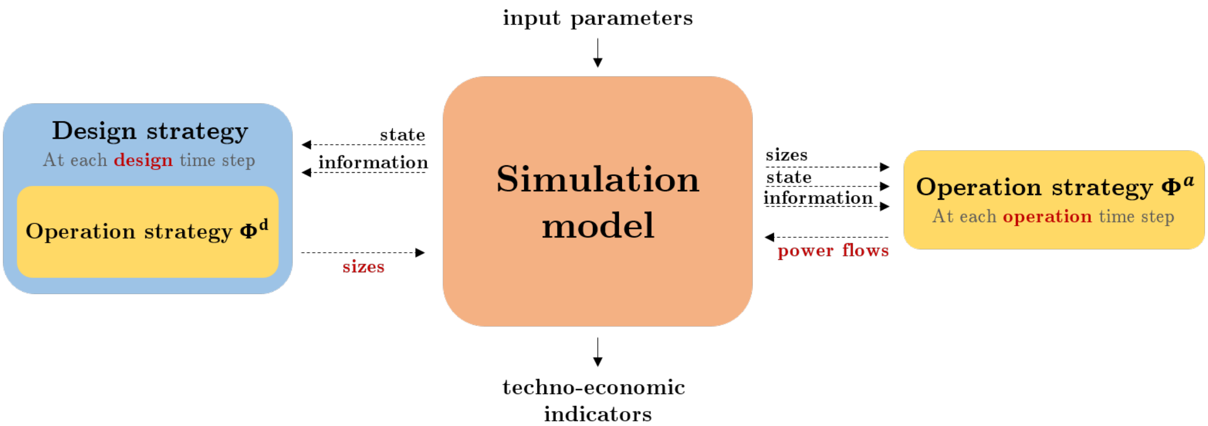


Figure 1.3: The techno-economic indicators of a project are computed by simulating both the operation and design strategies over a given time horizon. As depicted in the figure, the operation strategy used to design the system (Φ^d) can be different from the one used in real life (Φ^a).

[3]This is not always the case as decisions can be computed using heuristic rules.

[4]Optimization models are also based on techno-economic indicators. However, the indicators computed in the design phase are those of simplified problems. The "real" metrics of the project are those computed by the simulator.

The design process has a slow dynamic compared to the operation. Indeed, the adequacy between supply and demand must be guaranteed at all times while the decision to install or replace equipment is typically made on a slower time scale (e.g., years). Therefore, a distinction between the design and operation time steps is introduced to account for the different dynamics. The next paragraphs aim at describing more precisely both the operation and design strategies as they constitute critical aspects of this work.

Operation strategy

The objective of the operation strategy (also called *policy* or *energy management strategy*) is to determine the **power flow decisions** at each **operation time step** as a function of the available information (e.g., energy production and demand, the tariff of electricity) and state of the system (e.g., state-of-charge of the storage equipment). In this case, the size of the assets is fixed. Operation decisions are made to control the energy exchanged between the different equipment. The operating indicators are computed by simulating the operation of the system over a given time horizon.

Figure 1.4: The operation strategy gives, at each operation time step, the power flows as a function of the available information and state of the system.

Realistic strategies only have access to past and current information, meaning that they cannot anticipate the future. Decisions can be based on forecasts but these are computed from historical data. In a deterministic framework (i.e., all the information is perfectly known over the horizon), the ”perfect foresight” assumption means that the strategy has a perfect view of the future when making decisions. For example, the decision to charge the battery is made today because the policy knows, with perfect accuracy, that the production will be low tomorrow. This assumption is of course unrealistic but can be convenient to solve the design problem as explained in the next section. The operation strategy is named *anticipative* when it relies on this hypothesis in the following work.

The operation strategy can whether be based on a set of predefined rules or on optimization routines. Indeed, the problem is well-suited for optimization as it can be difficult to define a ”good” set of rules when the system complexity increases. Realistic optimized methods often rely on historical data, either to calibrate the algorithm (e.g., stochastic dynamic programming) or to make forecasts of uncertain parameters (e.g., model predictive control). The authors in [18] and, more recently, in [19] propose a classification of the different approaches to help the reader identify the specificities of each strategy. Finally, the optimized anticipative strategy gives the best (but unrealistic) solution over a given horizon. This information can be useful to evaluate the topology of the problem or to compare different operation strategies with each other. Further details are given in the next chapters.

Design strategy

The objective of the design strategy (also called *planning* strategy or *sizing strategy*) is to **select and properly size** the different assets at each **design time step** as a function of the available information (e.g., investment costs) and state of the system (e.g., remaining asset lifetimes). Note that the same previous remarks hold regarding the nature of the design strategies as the decision process is the same.

Since the design of the assets strongly depends on the system operation, planning models usually include the operation of the system over a given time horizon [14]. Hence, *integrated* design strategies refer to strategies where the operation of the system is integrated into the design formulation. In this case, the size of the assets is computed by "simulating" the operation strategy over a given time horizon (this is actually more complicated than it sounds because it depends on the approach used to solve the design problem. Chapter 2 provides additional details about this point).

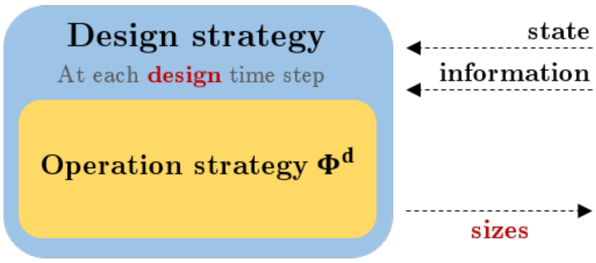


Figure 1.5: The design strategy gives, at each design time step, the size of the assets as a function of the available information and state of the system. Since the design of the assets strongly depends on the system operation, planning models usually embed the operation of the system over a given time horizon.

In the majority of design studies, the horizon is a single year [20, 21, 22, 23, 24]. Indeed, the design is computed based on the *equivalent annual cost* where the single year is considered as a representative period of the system lifetime. The result of the design strategy is a single sizing value which is computed at the beginning of the year. For this reason, the design problem is viewed as *static* (as opposed to *dynamic*) because a single decision needs to be made. In this case, the solution only gives a "snapshot" of the design at a given period, but the result does not provide any information about the long-term evolution of the system. On the other hand, the design can also be considered as a dynamic problem when the time horizon of the study is long enough to allow for multiple design decisions (equipment replacements due to aging for instance). In this case, simulating the design strategy provides, at the same time, the size of the assets and the investment pathway along the horizon (i.e., timing to install a technology).

In both cases, the integrated design problem is also suited for optimization as its resolution might be too complicated for a human brain [5]. The design problem is thus made of two nested optimization problems (i.e., design and operation) which must be solved jointly to theoretically guarantee optimality [25].

As further argued in the next section, the problem is often simplified for computational reasons. In the very specific case where the models used to simulate and design the DES are similar (e.g., same technological details, same operation strategy), the techno-economic indicators are the same in both phases. In this case, and only in this case, the simulator corresponds to the design model where the sizes have been fixed (only the operation is thus optimized). As explained in the next section, whether this choice is deliberate to emphasize a particular research point or it is unlikely to happen as simulation models are not suitable for optimization (because of computational tractability issues).

Because this work mainly deals with integrated design strategies, the following section aims at describing the different complexities attached to the problem and the typical simplifications made for optimization purposes.

1.2.2 A multifaceted problem

As previously mentioned, the integrated design of energy systems is a challenging task as many complexities are attached to the problem. Based on [14] and [17], figure 1.6 is a non-exhaustive and simplified view of the main complexities associated with the design problem. The figure is limited to the distributed energy system scope, and broader systems may include other ”macro-scale” aspects. For each category, the different ”levels” are given in increasing order of implementation complexity, according to the author’s point of view. They are described in the following.

[5]Methods based on extrapolation can be considered as heuristic strategies to solve simple design problems, and they have been used for years!

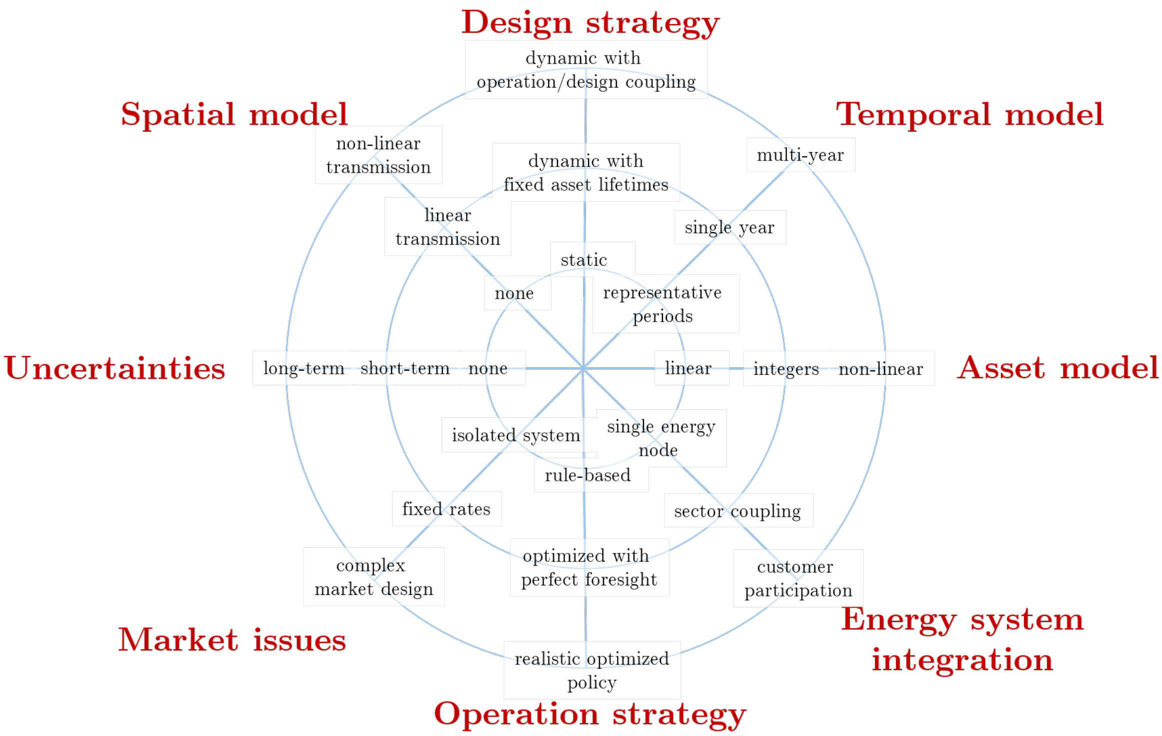


Figure 1.6: A non-exhaustive and simplified view of the different complexities associated with the integrated design problem. This graph is based on [14] and [17].

Temporal model

In a design problem, the temporal representation is a critical aspect as the number of variables (directly related to computational tractability) mainly results from the resolution and the horizon of the study. They are both crucial parameters as the former must be small enough to account for short time-scale variability, while the horizon should be long enough to fairly represent the long-term evolution of the system (e.g., system degradation, replacement due to aging). In the majority of the planning studies that include VRE, a one-hour time step is set to capture the variability of the production, while keeping reasonable computational times [14, 26]. Based on this operational constraint, several methods have been developed to reduce the time horizon in order to solve the design problem. A common approach is to select representative periods (e.g., representative days along the year) from the initial dataset. The impacts of such methods on the sizing results have been studied several times in the literature [15, 27, 28, 29, 30]. When the DES includes seasonal storage, yearly horizons are needed to cope with seasonal fluctuations. In this case, the authors in [23] and [31] found ingenious strategies to keep the problem tractable when modeling the full horizon is impossible. The highest accuracy level is reached when the full horizon is modeled over multiple decades with an hourly resolution, to account for long-term fluctuations. To the best of the author's knowledge, this problem is rarely addressed in the literature. Most of the studies dealing with the long-term evolution of energy systems whether use representative periods or coarser time resolution as mentioned in [32].

Spatial model

In our case study, ”spatial model” refers to the modeling complexity attached to the energy transmission networks between the equipment. It can whether be electrical power lines or/and any others transmission systems depending on the energy carriers. While transmissions are usually omitted in most of the DES planning studies (i.e., assuming system aggregation), this latter remark is not true for large-scale power systems [33, 34] where network related issues play an important role. The transmission complexities range from linear models to representations where non-linear phenomena are taken into account.

Uncertainties

Uncertainty is an important aspect of the problem as many parameters are unknown when sizing the system. They are usually classified into two categories: *aleatory* and *epistemic*. The former corresponds to the inherent stochastic nature of some phenomenon (e.g., the future solar production cannot be predicted with perfect accuracy even with a large number of measurements), and the latter is related to the knowledge that we have about the process (e.g., the model that we use might not be appropriate to represent the real problem). In-depth discussion about this classification can be found in [5, 35]. In this work, particular attention is paid to the information available when making decisions. For example, suppose that you have to decide today whether to invest in a battery. If you know with perfect accuracy that the cost will be halved tomorrow, you will probably wait for an extra day before buying the asset. On the other hand, if you have no information about how the cost will change, you might buy it today as it wouldn't make any difference to wait another day. Therefore, this simple example shows that decisions might be completely different whether or not the decision-maker has perfect foresight of the future. The role of information in the decision process has been broadly discussed in [19] for readers seeking more details about this important concern. In the following work, the problem is called *deterministic* when all the information is available over the horizon (i.e., uncertainties are not taken into account). Then, a difference is made between short- and long-term uncertainties. The former corresponds to the uncertainties attached to the operation parameters (e.g., the future energy demand and production are not perfectly known when making power flows decisions), while the latter concerns uncertainties related to the design parameters (e.g., future investment costs). Note that short-term uncertainties also affect the sizing results as the design depends on the system operation.

Asset model

Planning the design of energy systems involves choosing the right level of technical details to describe the equipment behavior. The modeling choice is always a trade-off between the physical model accuracy and the computational burden owning to the other facets of the problem. Does the technological aspect prevail over the other facets? The answer is not obvious *a priori* and there is no reason why a model which includes high technological details but completely omits the other facets such as uncertainty, for example, should give better design solutions. A common practice for the design of multi-energy systems is to use the *energy hub* concept first introduced by [36]. This modeling framework has been applied several times in the literature [16, 37] where energy conversion and storage systems are characterized through efficiencies [7]. A step further in the modeling accuracy is to use binary variables to linearize complex models when mathematical programming techniques are applied (e.g., see [38] and [39] for instance). This comes at a price as the combinatorial complexity increases, leading to longer computational times. A great review of the different modeling approaches is given by Allegrini et al [40] for readers seeking more details about the various options.

Energy system integration

In recent years, multi-energy strategies have gained popularity as they can provide flexibility to the system and facilitate the integration of VRE [8]. As a result, cross-sectoral studies are increasingly addressed in the literature, although power systems remain the most investigated [13]. Another important aspect is the integration of customer behaviors into the modeling frameworks. Examples include demand-side management issues or the response to incentives from multiple local emerging flexibility markets that may alter consumer energy demands (e.g., electric mobility). Most of the time, the energy demand is seen as exogenous information which is used as input of the planning model. However, the interaction with the customer can be integrated into the model by considering the energy demand as an elastic and endogenous parameter.

Market issues

In the majority of studies, the decision-maker seeks the least cost option while ensuring a given set of constraints. The total cost of the system is usually expressed as the sum of both the investment and operating costs. The simplest case (from a market perspective) is probably when the system is isolated, meaning that the total cost is limited to the investment and maintenance expenditures. On the other hand, when the DES can exchange energy with external networks, the project profitability directly depends on market rules, in addition to different economic incentive schemes [41, 42]. While fixed energy rates are easily addressed, the integration of new market designs is challenging as the DES owner can participate in multiple energy markets with various revenue streams.

Operation strategy

The different operation strategies have been described in the previous section. Most of the planning studies are based whether on anticipative or rule-based strategies because of computational issues [12, 14]. The reason for this is related to the nature of the methods used to solve the design problem. Further explanations are given in chapter 2 where different resolution approaches are also introduced. On the other hand, while optimized realistic operation strategies are widely studied in the literature [43], they are mainly approached without considering their relation with the design of the system: the authors usually start from a given DES where the equipment sizes are fixed, and the objective is to determine the operation strategy that gives the best performance.

Design strategy

Although the same classification on the nature of the design strategies would have been possible, the focus is set on whether the design is dynamic or not. As previously said, the problem is considered ”static” in most planning studies. Nevertheless, a growing body of literature is now emerging concerning the dynamic design of DES while accounting for detailed operation (see [44], [45] and [46] for instance). In this case, most of the studies assume that the assets have fixed lifetimes as it facilitates mathematical implementation. Indeed, because the asset lifetime is *a priori* known and does not depend on the operation, the decommissioning time is also known in advance. However, the interaction between the design and the operation is not fully captured as the way systems are operated has no consequences on their aging. For example, this latter point can be particularly critical for a battery because the number of charge/discharge cycles has great consequences on the equipment’s lifetime. Including the impact of the operation on the design in planning models is challenging, and thus rarely addressed in the literature. This issue is addressed in chapter 5 of this thesis.

Since the main complexities of the design problem have been described, the objective of the following section is to identify some directions for further research in order to define the scope and objectives of this thesis.

1.2.3 The need for further research

Before going into details, a first remark must be made: no energy planning model includes all features with the highest level of accuracy in the literature! This is due to obvious computational reasons. Any planning model is an approximation of the real problem and simplifications are made according to the objectives of the study it serves. The modeling art is to carefully determine which simplification is consistent with the purpose of the study [6]. This point is actually good news for us because it leaves room for research!

Having said that, some directions are more investigated than others when dealing with the design of DES. However, it might be difficult to draw a "well-known" region on the radar plot above (figure 1.6) because it is more about combinations of the different levels than a level in itself. For example, it does not mean anything to say that "sector coupling has been addressed several times in the literature" because if this feature is associated with non-linear asset and transmission models along with long-term investment uncertainties, then the problem remains unsolved. Therefore, the number of unexplored combinations must be considerable and it would make no sense to try to list them all. Instead, critical ongoing research challenges are identified from recent journal reviews and thesis (at least published in 2019). Examples include Chang et al [13] (which can be seen as an updated version of Pfenninger et al [48]), Cuisinier et al [14], Mavromatidis et al [16], Helisto et al [17] and Sepulveda [44].

The first set of challenges concerns the mathematical methods to solve the optimization models. Endless efforts are needed to reduce computation time as it makes it possible to include more details in the models, leading to a better representation of reality. For instance, these involve improving decomposition techniques (or mixing them with each other), introducing high-performance computing that usually requires rethinking the model architecture, and assessing the potential of machine learning approaches to speed-up computations. Meeting these challenges requires a strong mathematical background and mathematicians are the natural target for these tasks.

Another direction pertains to the integration of uncertainties at both time scales while keeping a high resolution for the operation. Indeed, uncertainties are usually omitted in the majority of studies because the complexity rapidly increases, leading to intractable problems. However, there is a general agreement that planning models must contend with uncertainties as they can strongly affect the results. This point is directly related to another aspect that concerns the "perfect foresight" hypothesis. Indeed, realistic design and operation strategies do not have access to future values of the uncertain parameters (e.g., energy demands and production, the future tariff of electricity and investment costs). They can use forecasts to make decisions, but the exact future values are unknown. However, most of the planning tools rely on this assumption without assessing the true performance of the resulting design. For example, there is no reason *a priori* that the requirements will be met with a realistic operation strategy if the one used to design the system relies on the perfect foresight hypothesis. This latter assumption is a modeling approximation to facilitate the resolution of the design optimization problem, but its true relevance must be (at least!) discussed regarding real-life operation. This problem will

[6]See for example [47] for a thorough discussion about common simplifications made for optimization purposes

be tackled in chapter 4 of this thesis.

This challenge raises a more general question about the simplifications made to solve the design issue. Indeed, modeling for energy systems is always a trade-off between computation time and model accuracy. However, it is not clear what should be simplified without compromising the results and how critical one simplification is compared to another. The answer is not obvious *a priori* and might depend on the objective of the case study. For example, several studies have already investigated the impact of technological details [39, 49] or temporal representation [27, 29] on design outcomes, but the same evaluation should also be done jointly for the other facets of the problem. Is the "perfect foresight" simplification more critical than using constant efficiencies to describe the asset behavior? Such comparative studies can provide valuable insight into these issues.

Another important direction for long-term studies is to consider multi-year dynamic designs with great operational details while accounting for design and operation uncertainties. Unlike "snapshot" investment, these models also provide the investment pathway which can be valuable information for policy and decision-makers. The integration of this feature into planning models significantly increases the complexity of their resolution, especially if the precision of the other facets is kept. As mentioned earlier, the interaction of the operation on the design dynamic seems to be partially addressed in the literature while it may be a relevant issue for storage systems. This latter issue will be the focus of chapter 5.

In addition to the previous challenges, another aspect that is frequently mentioned is the need to address cross-sectoral analysis with endogenous demands to include customer behavior in the modeling framework. This latter point is less studied in the literature but seems to be a critical aspect to account for uncertain changes in demand and its interaction with the system operation.

Finally, the last challenging issue described in this section concerns the connection of these tools to the real world. This point is obviously out of the scope of this thesis, but it represents a critical issue for the relevance and reliability of these planning approaches. Indeed, the objective of these methods is to support stakeholder and policy-maker decisions, so energy models must be accessible and understandable. Many efforts are probably needed to improve their transparency and to encourage collaboration between modelers and other actors to ensure that the modeling tendencies are in-line with end-user needs.

The previous challenges represent a tiny fraction of the potential unexplored directions, but they are more than sufficient to determine some research orientations for this thesis. Based on these limitations, the next section introduces the scope and objective of this work.

1.3 Scope and objectives

Based on these challenging issues, the possible directions of this work are unlimited. But the PhD has a finite and short horizon, especially for someone who is barely familiar with the energy planning realm.

Therefore, this thesis revolves around four research items presented in section 1.3.1. They constitute the line of sight of this work (it does not mean that these questions will be resolved in the end, but they motivate the following developments). Section 1.3.2 describes and justifies the modeling choices made to address these challenges.

1.3.1 Questions addressed in this work

This work basically revolves around four key points given as follows:

1. What is the additional value brought by **multi-energy systems and seasonal storage** to supply customers with a high share of solar production? In particular, what is the contribution of the **hydrogen cogenerated heat**? These aspects will be discussed in chapter 2.

2. How can **uncertainties** be included in energy planning studies? What kind of information is available (at each time step) to make both operation and design decisions? How does it impact the results? These points will be addressed in chapter 4.

3. What about the **interaction between the design and operation** strategies? Should the operation strategy used to design the system and the one used in real-time be identical? These questions are also discussed in chapter 4.

4. How can the **impact of the operation on equipment aging** be included in **dynamic planning** models? Does it bring more value compared to design strategies based on "snapshot" investment? Chapter 5 will be dedicated to these questions.

An additional practical issue arises from the previous points as different combinations of design and operation methods will be compared in the following work. However, to the best of the author's knowledge, there is no existing tool that can easily accommodate multiple sizing and operation strategies without changing the overall code. This point is further discussed in chapter 6. Therefore, the practical coding issue is:

5. How can a **generic tool** be implemented to **compare and evaluate multiple combinations of design and operation strategies**?

1.3.2 Modeling choices

A model is a mathematical representation of real processes. By nature, it is a simplified and truncated view of reality to facilitate its apprehension by humans. Thus, a model inevitably comes with a large set of assumptions that must be clearly stated before using it. In this way, a model **does not** tell the ground truth, but it can be a useful tool for decision support as long as the main features of the problem are correctly described. Therefore, the model granularity must be chosen according to the research purpose. While quantum mechanics is not appropriate to deal with electrotechnical issues (even if they both deal with electron phenomena), the level of detail attached to planning models must be questioned in the same consistent manner. In addition to the research objectives, another aspect that should motivate the modeling choice is its adequacy with the models mainly used in the research community. Indeed, this work is part of a more global energy modeling environment that already has its own modeling practices. Hence, the model granularity should also be consistent with the literature to allow comparison of the results and fast integration with other modeling activities.

Before going into more details, an important remark concerns the validation of such planning models based on optimization. While it is relatively easy to validate simulation models, it has been argued several times that the validation of design models is a more complicated task [48, 50]. Indeed, the objective of design models is to find out the optimal system configuration according to long-term techno-economic objectives and constraints. As stated by Mavromatidis et al in [16]: "*it would be very hard or even impossible to check whether the optimal configuration for a DMES predicted by a design optimization model is indeed optimal with real-life experiments, as such an investigation would require installing multiple energy systems and comparing their long-term performance.*". Therefore, this latter point does not mean that planning models are useless (it might be better than nothing), but it reinforces that they should be carefully used, and results strongly criticized. **Design optimization models should be considered as additional pieces to support decisions, but they cannot constitute the only basis**.

Before presenting the assumptions made in this work, note that each of the ingredients previously introduced (i.e., design strategy, operation strategy and the simulator) is a research issue in itself. To rigorously address the overall problem and yield reliable quantitative answers, the three directions must be thoroughly conducted in parallel. This would obviously require more than 3 years (at least for me!). Therefore, **this thesis mainly focuses on methodological aspects attached to design strategies**. However, as explained above, the design issue cannot be treated without any connection to the operation and simulation of the system. In the following, the only differences between the models used to simulate and design the DES are introduced to emphasize some specific research aspects, thoroughly described below. In addition, case studies and technological details are kept deliberately simple as the objective is to illustrate design and operation issues rather than giving extensive quantitative results. As shown in the following chapters, even with simple models, the problem remains difficult and many questions are still unsolved.

Having said that, the following paragraphs give an overview of the modeling choices made in this work. For each research question, the modeling assumptions are depicted on the radar chart from section 1.2.2 to get a clear picture of what lies behind the models that are implemented.

Therefore, to address the first research question in the next chapter, the multi-energy assets and the seasonal storage have low technological details. The model is derived in a deterministic framework over a single equivalent year because the objective is to obtain a problem that can be solved quickly. This allows running multiple parametric analyses to show global trends and foster a systemic understanding of how technologies interact. In this case, there is no difference between the simulator and design models. The modeling assumptions are depicted in figure 1.7.

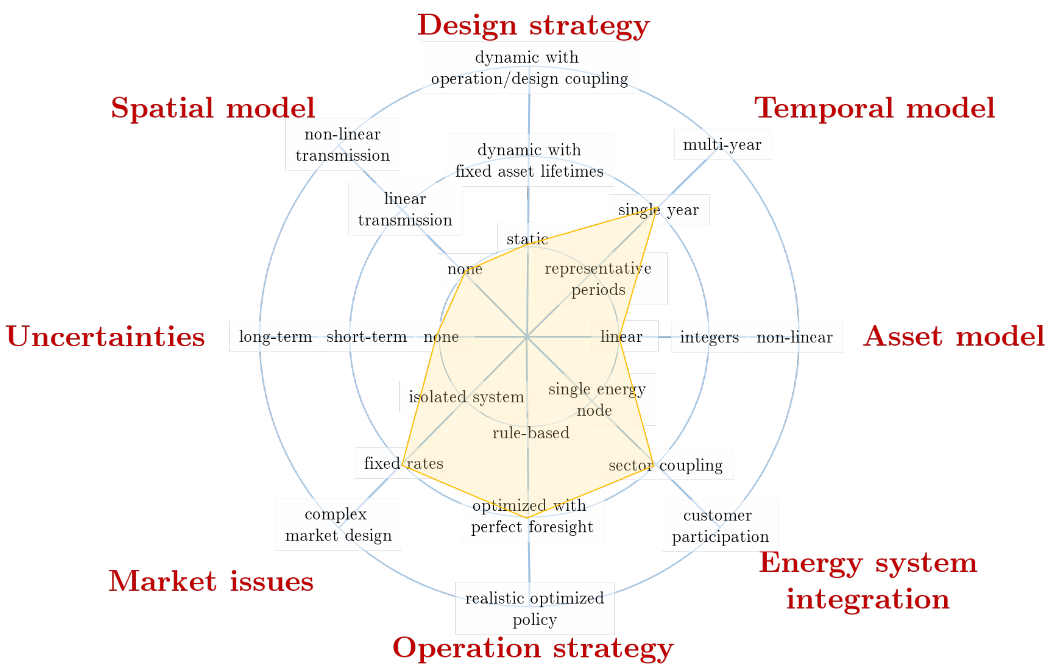


Figure 1.7: Modeling assumptions made to tackle the first part of this work (see chapter 2).

The second and third questions (addressed in chapter 4) deal with stochastic issues and the interactions between both the design and operation strategies. Therefore, the stochastic model is built from the deterministic model previously introduced. Realistic optimized policies and operation uncertainties are added to the framework as shown in figure 1.8. Special attention is paid to the information available to make both operation and design decisions. To this end, the only difference between the optimization and simulation phases is the way information is revealed over time: the design strategy may rely on the "perfect foresight" assumption, while the simulation is run with realistic operation strategies that only have access to past and current information.

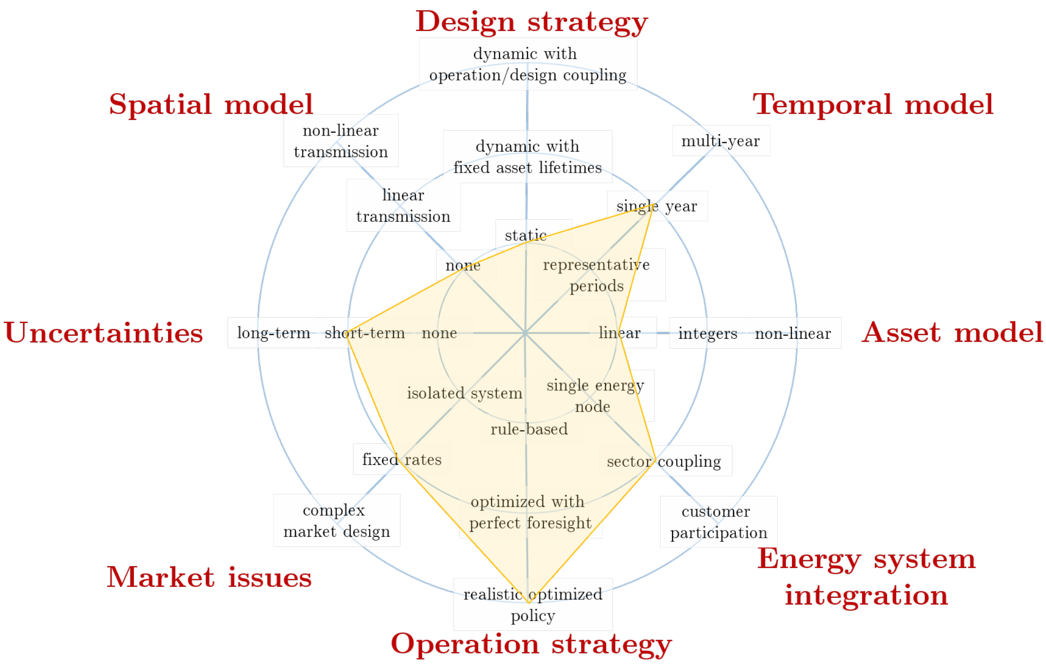


Figure 1.8: Modeling assumptions made to tackle the second part of this work (see chapter 4).

The last question (addressed in chapter 5) concerns the dynamic design of DES which includes the impact of the operation over system lifetimes. As it significantly complicates the problem resolution, the DES only includes solar panels and a battery storage in a deterministic framework. As this question is rarely addressed in the literature, the objective is to focus on the design methodologies rather than the system complexity. Another goal is also to compare the aware aging method to more common design approaches. To this end, the simulator includes the dynamic of the design properly modeled, while design strategies might be based on different modeling choices. The modeling assumptions are depicted in figure 1.9.

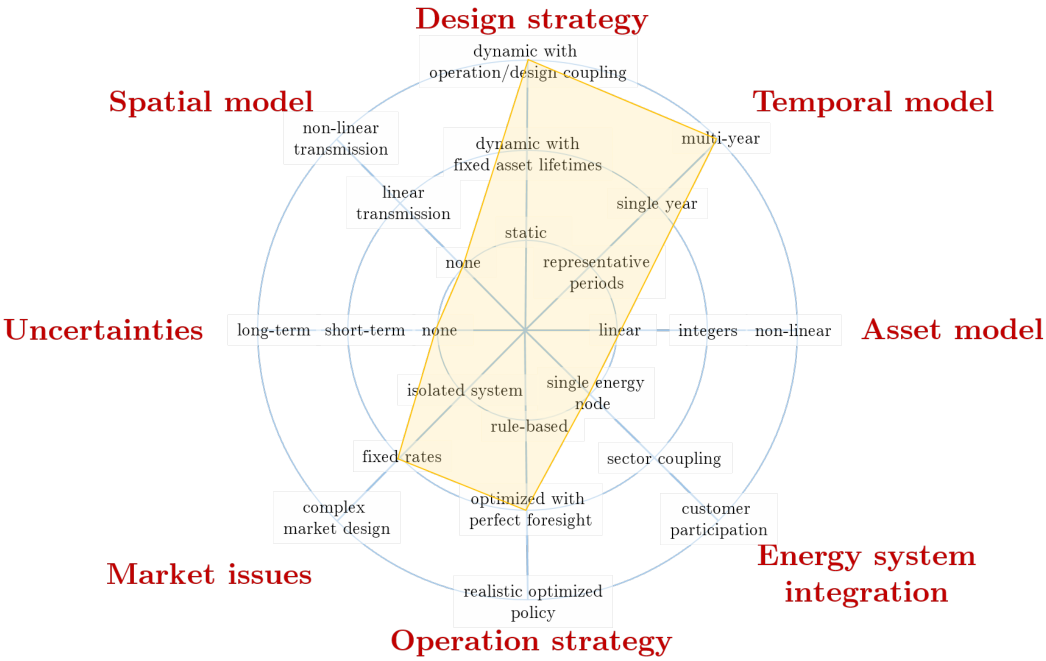


Figure 1.9: Modeling assumptions made to tackle the third part of this work (see chapter 5).

1.4 Thesis overview

This work is structured into seven chapters where the research questions that were presented above are chronologically addressed. In addition to the multi-energy system benchmark, a *toy-example* is also introduced in the following. The toy-example remains simple as it only includes solar panels and a battery storage. The objective is to help newcomers understand the planning problem by providing the entire formulation on a simple case study without the complexity brought by the multi-energy aspects. In each chapter, the overall formulation of the toy-example is given in a colored box and anyone only interested in the equations can directly refer to this box without scanning the whole document (see figure 1.10)

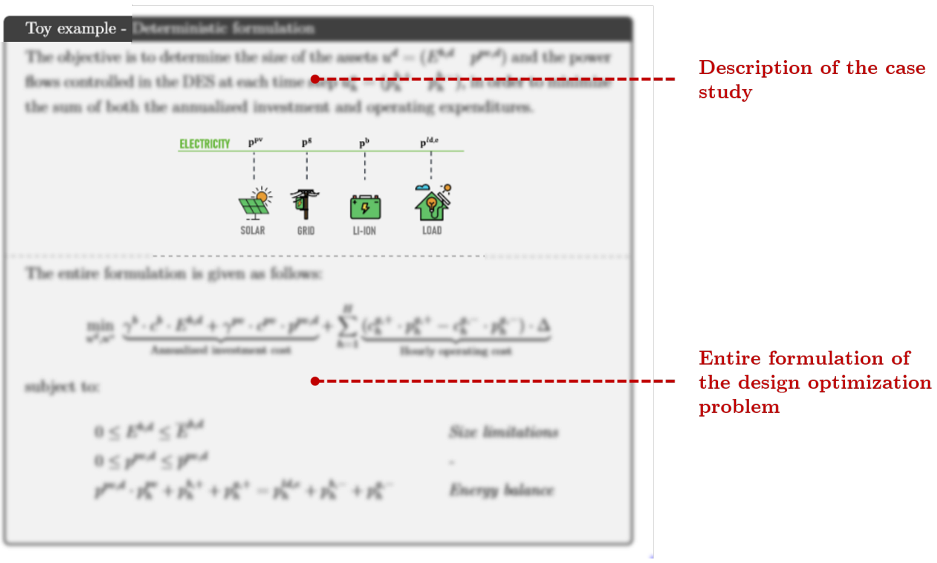


Figure 1.10: The overall formulation of the toy-example design problem is given in a colored box and anyone only interested in the equations can directly refer to this box without scanning the whole document.

Chapter 2 deals with the deterministic design of a multi-energy system with seasonal storage to supply residential customers with a high share of solar production. The objective of this part is two-fold: 1) introduce the deterministic model that will later form the basis of the stochastic approaches; 2) identify a case study where a multi-energy system with seasonal storage is valuable. The contribution of this chapter to the state of the art is limited but this part is an essential prerequisite to properly understand the next steps.

Chapter 3 presents a method based on Markov Chains to generate a large number of energy demands and production scenarios. These are necessary ingredients for the stochastic approaches developed in the next chapter.

Chapter 4 addresses the design of DES under uncertainties and discusses the impact of the operation strategy in the design procedure. To this end, two design methods based on stochastic programming are developed (i.e., one is based on mathematical programming and the other uses a metaheuristic algorithm). These approaches are based on the formulation of chapter 2. The sizing solutions are then compared using several realistic operation policies (i.e., rule-based and look-ahead methods) which only have access to past and current information. This chapter discusses the interplay between the design and the operation by jointly evaluating and comparing the different approaches.

Chapter 5 deals with the dynamic design of DES with the specificity previously introduced. In this case, the model takes technology replacement due to aging into account, so multiple design decisions have to be made over the horizon. The dynamic aware aging design method (which comes from the literature) is then compared to two heuristic design strategies based on single representative years. Note that this study was carried out before chapters 3 and 4, but it feels to the author that the thesis is more coherent structured in this way.

Chapter 6 presents the architecture of the Genesys.jl toolbox that has been developed throughout this thesis. This toolbox enables to easily assess and compare different design and operation strategies without changing the overall code. This part also shows how custom approaches can be easily added to the framework.

Finally, **Chapter 7** summarises the main conclusions of this thesis and gives future research directions for this work.

Part I

A deterministic framework for the design of multi-energy systems

Chapter 2

The value of multi-energy systems and seasonal storage to supply residential customers with a high share of solar production

Highlights

- Introduction of the deterministic model that will form the basis of the stochastic formulation.
- Parametric analysis of the renewable share and electricity tariff over the design of the DES.
- The total annual cost increases exponentially with the renewable share and the integration of a low-cost dispatchable source may significantly alleviate the overall system cost.
- To reach autonomy, the multi-energy system with seasonal storage shows better techno-economic performance than the battery-only solution. Cogenerated heat from hydrogen helps reduce the cost.

Contents

2.1 Introduction

The first step of this thesis is to introduce the deterministic model that will serve as a basis for the next chapters. In this deterministic case, all the input parameters (i.e., future energy demands, production and electricity tariffs) are known with certainty over the entire horizon. As previously said in the introduction, the contribution of this part to the state of the art (from a methodological point of view) is limited but this chapter is an essential prerequisite to properly understand the next steps. Another goal is also to provide insights into the value of multi-energy systems and seasonal storage to supply residential customers with a high share of solar production and to get a systemic understanding of how technologies interact.

Deterministic optimization models have been widely used in the literature for the design of DMES [12]. For example, Gabrielli et al [23] used a deterministic model to design a multi-energy system with seasonal storage. Especially, they introduced a method to reduce computation times by simplifying the temporal representation of the initial problem. Stadler et al [51] implemented such a model to optimize the size of distributed and renewable energy systems in buildings. Yang et al [52] did the same exercise at a district scale where another objective is to find the optimal location for the equipment. Murray et al [53] applied this approach to assess the potential of long- and short-term storage systems in decentralized neighborhoods. The model was solved for different scenarios to analyze potential future developments of the technologies. Finally, Cuisinier et al [14] provided a review of optimization models for the planning of local energy systems where deterministic models are in the majority. In all the previous references, the model is formulated on a single representative year where the investment costs have been annualized according to the equipment lifetimes. Based on the literature, the reasons that usually motivate deterministic modeling are mostly the following: 1) technologies are represented with high levels of details, thus including uncertainties would lead to intractable problems; 2) the authors need models that can be quickly executed to run parametric analysis (see for instance [5, 54, 55]).

To solve these problems, two categories of methods are mainly represented in the literature: **mathematical programming** and **methaheuristics** [5, 12]. In the former case, a single large model based on mathematical programming is formulated where both the design and operation decisions are variables of the same problem. Among the different models, linear programming (LP) or mixed-integer linear programming (MILP) problems are the dominant approaches [12]. In the latter case, binary variables are usually introduced to linearize complex models using piecewise linear approximation [56]. Decomposition methods (e.g., Bender decomposition [57]) are often needed when large-scale problems are difficult to solve. These methods usually rely on external solvers (e.g., Cbc [58], CPLEX [59]) to find the solutions. The main drawback of these approaches is that the resolution complexity increases rapidly with binary variables. Therefore, either the models remain linear to avoid binaries or other simplifications are generally needed (e.g., representative periods introduced in chapter 1) to solve the problem.

On the other hand, metaheuristic approaches use black-box algorithms to compute the size of the assets. In this case, the first step is to build a simulator of the system operation, unless it comes from an external source. Then, the simulator is embedded in an external loop that investigates different system configurations via a metaheuristic algorithm. Examples of such algorithms include Particle Swarm Optimization (PSO) [60], Genetic Algorithm (GA) [24] and niching methods [61] to name just a few. The optimal solution is not guaranteed but if the algorithm is properly configured, this approach is effective in finding the optimum (or at least, getting close to it). The main advantage is that the algorithm can solve highly nonlinear problems that allow representing the behavior of technologies with physical models closer to reality. However, the simulation of the operation strategy must be fast because the design loop requires multiple iterations to converge towards the solution. Most of the time, the operation strategy is either based on heuristic rules [62, 63] or anticipative policies [24, 64].

In recent years, another technique has gained popularity where both optimization and simulation tools are jointly used in an iterative fashion (see for instance [65, 66, 67, 68]). This approach is named *soft-linking*. The idea behind this is to get "the best of the two worlds" by using efficient optimization and simulation models to converge toward a feasible solution. Indeed, highly detailed models cannot be used for optimization because of computational tractability. Thus, the design is computed with a simplified version of the problem and the sizes are evaluated on a simulator which is a good representation of the real problem. Then, a design optimization parameter is progressively tuned by successive iterations between the two phases until the requirements are met.

In the following work, the objective is to perform multiple parametric analyses, meaning that the model must be solved quickly. Furthermore, the integration of uncertainties in the next chapters will significantly complicate the problem resolution, thus the level of details remains simple as explained in chapter 1. Therefore, a large LP model is formulated over a single representative year for the design of the DMES depicted in figure 2.1. The main objective of the multi-energy system is to supply both residential electrical and thermal demands while ensuring a given share of solar production (also called "renewable share"). This ratio represents the share of the total consumption supplied by the local solar production. The multi-energy system is connected to the electrical grid and may include a set of energy converters and storage systems. The results will be compared to a reference case where all the energy is purchased from the utility grid. In this baseline case, the thermal demand is supplied thanks to a heater that converts electricity to heat. The main goal of this study is to assess how the design is changing regarding different values of renewable share and electricity tariff.

First, the mathematical formulation is shown in section 2.2. Then, the problem is solved regarding several values of renewable share and electricity tariff in section 2.3. Finally, the results are discussed in section 2.4, followed by the conclusion.

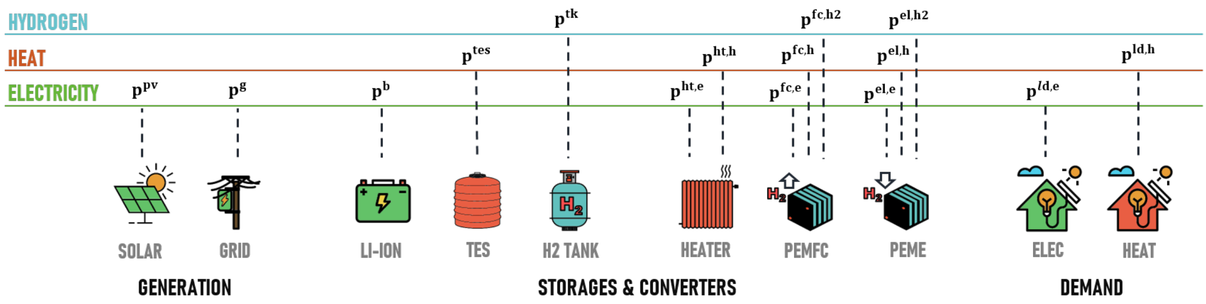


Figure 2.1: Schematic view of the multi-energy system with seasonal storage.

2.2 Mathematical formulation

This section shows the mathematical formulations of the deterministic optimal design and operation problems.

2.2.1 Decision variables

The decision variables for both the design and operation optimization problems are given as follows:

- The decision variables for the design are the size of the assets. They correspond to the maximum capacity of the storage systems and the maximum power of the energy converters, gathered in the vector $u^d \in \mathbb{U}^d = \mathbb{R}^6$ (2.1). Note that the size of the heater is not a decision variable of the problem because a 10 kW heater is assumed to be already installed for the baseline case.

$$u^d = (E^{b,d} \quad E^{tes,d} \quad E^{tk,d} \quad p^{fc,d} \quad p^{el,d} \quad p^{pv,d}) \tag{2.1}$$

- The decision variables for the operation are the power flows controlled in the multi-energy system at every time step. They correspond to the charging and discharging power for the storage systems and the electrical powers for the energy converters, gathered in the vector $u_h^o \in \mathbb{U}_h^o = \mathbb{R}^9$ (2.2).

$$u_h^o = (p_h^{b,+} \quad p_h^{b,-} \quad p_h^{tes,+} \quad p_h^{tes,-} \quad p_h^{tk,+} \quad p_h^{tk,-} \quad p_h^{fc,e} \quad p_h^{el,e} \quad p_h^{ht,e}) \tag{2.2}$$

2.2.2 Constraints

The set of constraints defines the admissible set of solutions for both the design and operation problems.

Maximum equipment sizes

The size of the assets is limited by physical constraints set by the decision-maker. For example, the PV peak power is directly bounded by the available surface required to install solar panels. These limitations materialize by equations (2.3)-(2.8).

$$0 \leq E^{b,d} \leq \overline{E}^{b,d} \tag{2.3}$$
$$0 \leq E^{tes,d} \leq \overline{E}^{tes,d} \tag{2.4}$$
$$0 \leq E^{tk,d} \leq \overline{E}^{tk,d} \tag{2.5}$$
$$0 \leq p^{fc,d} \leq \overline{p}^{fc,d} \tag{2.6}$$
$$0 \leq p^{el,d} \leq \overline{p}^{el,d} \tag{2.7}$$
$$0 \leq p^{pv,d} \leq \overline{p}^{pv,d} \tag{2.8}$$

Energy balances

At each time step, operation decisions are made in order to ensure the energy balance between production and demand for every energy node (i.e., electricity, heat and hydrogen respectively). It materializes by equations (2.9)-(2.11).

$$p^{pv,d} \cdot p_h^{pv} + p_h^{b,+} + p_h^{fc,e} + p_h^{g,+} = p_h^{ld,e} + p_h^{b,-} + p_h^{el,e} + p_h^{ht,e} + p_h^{g,-} \tag{2.9}$$
$$p_h^{tes,+} + p_h^{el,h} + p_h^{fc,h} + p_h^{ht,h} \geq p_h^{ld,h} + p_h^{tes,-} \tag{2.10}$$
$$p_h^{tk,+} + p_h^{fc,h2} = p_h^{tk,-} + p_h^{el,h2} \tag{2.11}$$

where p_h^{pv} is the hourly solar capacity factor between 0 and 1, while $p_h^{ld,e}$ and $p_h^{ld,h}$ are the electrical and thermal inelastic demand, respectively. The electrical grid powers $p_h^{g,+}$ and $p_h^{g,-}$ are introduced to make the optimization implementation clearer but they are not decision variables of the problem. Their values are computed at the end of each time step and limited by the maximum power allowed by the external network $\overline{g}$ (2.12)-(2.13).

$$0 \leq p_h^{g,+} \leq \overline{g} \tag{2.12}$$

$$0 \leq p_h^{g,-} \leq \overline{g} \tag{2.13}$$

Note that no such variable exists for the thermal demand in our model: a lack of supply would result in temperature discomfort. Also, the thermal energy balance (2.10) is given with inequality as the overproduced heat is assumed to be dissipated in a dump load. This simple trick avoids the introduction of another decision variable for the operation as it has no consequences on the results.

Energy system model

The physical models of the assets remain deliberately simple as the objective of this work is to focus on parametric analysis, rather than the physical model complexity. Thus, the modeling is based on the energy hub concept from [69] which is widely used in the literature when dealing with the design of distributed multi-energy systems. In-depth discussions about modeling implications and related issues could be found in [16] and [48] for instance.

Storage systems. Three types of storage technologies are considered in the multi-energy system: (i) a Li-ion battery, (ii) a thermal energy storage (TES) and (iii) a high-pressure hydrogen tank. A generic storage model is implemented for both equipment and the state of charge dynamic is given by equation (2.14).

$$E_{h+1} = E_h \cdot (1 - \eta^{loss} \cdot \Delta h) + (\eta^- \cdot p_h^- - \frac{p_h^+}{\eta^+}) \cdot \Delta h \tag{2.14}$$

$$\underline{e} \cdot E^d \leq E_h \leq \overline{e} \cdot E^d \tag{2.15}$$

$$E_1 \leq E_{H+1} \tag{2.16}$$

where E_h is the state of charge expressed in kWh, η^- and η^+ are respectively, the charging and discharging efficiencies, η^{loss} is the self-discharge coefficient and Δh the operation time step (in hours). The state of charge is bounded by a percentage of the maximum storage capacity E^d (2.15). A periodicity constraint (2.16) is also added to the storage state of charge. p_h^- and p_h^+ are the charging and discharging powers which are limited by equations (2.17) and (2.18) as the maximum amount of energy that could be exchanged during a time step is limited.

$$0 \leq p_h^+ \leq \overline{p} \cdot E^d \tag{2.17}$$

$$0 \leq p_h^- \leq \underline{p} \cdot E^d \tag{2.18}$$

The state variables of the problem are gathered in the vector $x_h \in \mathbb{X}_h = \mathbb{R}^3$ (2.19).

$$x_h = (E_h^b \quad E_h^{tes} \quad E_h^{tk}) \tag{2.19}$$

Energy converters. They transform a given energy carrier into another through specific processes, attached with energy efficiencies. The heater produces heat thanks to electricity and the coupling equation is given by (2.20). The proton-exchange membrane fuel cell (PEMFC) produces electricity, heat and water from hydrogen (2.21)-(2.22). The chemical reaction is exothermic and the heat produced could be extracted to supply a thermal demand. On the contrary, the proton-exchange membrane electrolyzer (PEME) produces hydrogen from electricity and water (2.23)-(2.24). Overproduced heat by losses inside the component could also be extracted.

$$p_h^{ht,h} = \eta^{ht,e\to h} \cdot p_h^{ht,e} \tag{2.20}$$

$$p_h^{fc,e} = \eta^{fc,h_2\to e} \cdot p_h^{fc,h2} \tag{2.21}$$

$$p_h^{fc,h} = \eta^{fc,h_2\to h} \cdot p_h^{fc,h2} \tag{2.22}$$

$$p_h^{el,h2} = \eta^{el,e\to h2} \cdot p_h^{el,e} \tag{2.23}$$

$$p_h^{el,h} = \eta^{el,e\to h} \cdot p_h^{el,e} \tag{2.24}$$

where $\eta^{ht,e\to h}$ is the heater conversion efficiency from electricity to heat. $\eta^{fc,h_2\to e}$ and $\eta^{fc,h_2\to h}$ are the PEMFC conversion efficiencies from hydrogen to electricity and heat, respectively. $\eta^{el,e\to h2}$ and $\eta^{el,e\to h}$ are the PEME conversion efficiencies from electricity to hydrogen and heat, respectively.

The converter electrical powers $p_h^{ht,e}$, $p_h^{fc,e}$ and $p_h^{el,e}$ are positive and limited by the size of the assets (2.25)-(2.27).

$$0 \leq p_h^{ht,e} \leq p^{ht,d} \tag{2.25}$$

$$0 \leq p_h^{fc,e} \leq p^{fc,d} \tag{2.26}$$

$$0 \leq p_h^{el,e} \leq p^{el,d} \tag{2.27}$$

Share of solar production

The share of solar production $\tau^{sh} \in [0,1]$ represents the proportion of the total consumption supplied by the local solar production: a ratio equal to 1 means that all the electricity is provided on-site. On the other hand, a ratio equal to 0 means that the total consumption (i.e., electrical and thermal demands through the heater efficiency) is provided by the utility grid. Thus, its value is computed based on the energy imported from the grid and the baseline total consumption (2.28).

$$\sum_{h=1}^{H} \Big(p_h^{g,+} - (1 - \tau^{sh}) \cdot (p_h^{ld,e} + \frac{p_h^{ld,h}}{\eta^{e \to h}}) \Big) \cdot \Delta h \leq 0 \tag{2.28}$$

2.2.3 Optimization problem statement

The integrated design objective is to determine the sizing and operation decisions in order to minimize the sum of both the annualized investment and operating expenditures as it is commonly done in the literature [20, 21, 22, 23, 24].

Annualized investment cost

The annual capital cost depends on the equipment sizes which are the investment decisions of the design problem u^d, and the capital cost of each technology (2.29).

$$J^d(u^d) = \gamma^b \cdot c^b \cdot E^{b,d} + \gamma^{tes} \cdot c^{tes} \cdot E^{tes,d} + \gamma^{tk} \cdot c^{tk} \cdot E^{tk,d} + \gamma^{fc} \cdot c^{fc} \cdot p^{fc,d} + \gamma^{el} \cdot c^{el} \cdot p^{el,d} + \gamma^{pv} \cdot c^{pv} \cdot p^{pv,d} \tag{2.29}$$

where the standard annuity factor γ (2.30) is computed based on the expected lifetime L of each technology (in years) with an interest rate r of 4.5 %.

$$\gamma = \frac{r \cdot (r+1)^L}{(r+1)^L - 1} \tag{2.30}$$

Operating cost

At each time step, the operating cost is computed based on the energy exchanged with the utility grid (2.31).

$$J_h^o(u^d, u_h^o) = (c_h^{g,+} \cdot p_h^{g,+} - c_h^{g,-} \cdot p_h^{g,-}) \cdot \Delta h \tag{2.31}$$

where $c^{g,+}$ is the tariff of electricity (€/kWh) and $c^{g,-}$ the feed-in tariff (€/kWh). Note that for the sake of simplicity, the maintenance costs are not included in this work but they can be easily added (for instance, as a fraction of the annual capital cost following [23]).

Problem statement

The objective of the design optimization problem is to determine the design and operation decisions in order to minimize the equivalent annual cost. The problem statement is given by equations (2.32).

$$\min_{u^d, u^o} \; J^d(u^d) + \sum_{h=1}^{H} J^o_h(u^d, u^o_h) \tag{2.32a}$$

s.t.

$$x_{h+1} = f(x_h, u^d, u^o_h) \tag{2.32b}$$

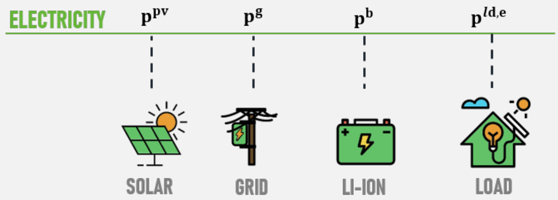
$$u^d \in U^d, \quad u^o_h \in U^o_h(u^d, x_h) \tag{2.32c}$$

where f is described by the state equations (2.14) for storage systems. U^d and U^o_h are the set of admissible solutions for the design and the operation, defined by the constraints introduced in section 2.2.2.

The resulting optimization problem belongs to the *linear programming* (LP) category, which can be easily solved with traditional solvers like CPLEX [59].

Toy example - Deterministic formulation

The objective is to determine the size of the assets $u^d = (E^{b,d} \quad p^{pv,d})$ and the power flows controlled in the DES at each time step $u^o_h = (p^{b,+}_h \quad p^{b,-}_h)$, in order to minimize the sum of both the annualized investment and operating expenditures. Remind that the electrical grid powers $p^{g,+}_h$ and $p^{g,-}_h$ are introduced to make the optimization implementation clearer but they are not decision variables of the problem.



The entire formulation is given as follows:

$$\min_{u^d,u^o} \underbrace{\gamma^b \cdot c^b \cdot E^{b,d} + \gamma^{pv} \cdot c^{pv} \cdot p^{pv,d}}_{\text{Annualized investment cost}} + \sum_{h=1}^{H} \underbrace{(c^{g,+}_h \cdot p^{g,+}_h - c^{g,-}_h \cdot p^{g,-}_h) \cdot \Delta h}_{\text{Hourly operating cost}}$$

subject to:

$$\begin{aligned}
&0 \leq E^{b,d} \leq \overline{E}^{b,d} && \textit{Size limitations}\\
&0 \leq p^{pv,d} \leq \overline{p}^{pv,d} && \text{-}\\
&p^{pv,d} \cdot p^{pv}_h + p^{b,+}_h + p^{g,+}_h = p^{ld,e}_h + p^{b,-}_h + p^{g,-}_h && \textit{Energy balance}\\
&0 \leq p^{g,+}_h \leq \overline{g} && \textit{Power grid}\\
&0 \leq p^{g,-}_h \leq \overline{g} && \text{-}\\
&E^b_{h+1} = E^b_h \cdot (1 - \eta^{loss} \cdot \Delta h) + (\eta^- \cdot p^{b,-}_h - \frac{p^{b,+}_h}{\eta^+}) \cdot \Delta h && \textit{Storage model}\\
&\underline{e} \cdot E^{b,d} \leq E^b_h \leq \overline{e} \cdot E^{b,d} && \text{-}\\
&0 \leq p^{b,+}_h \leq \overline{p} \cdot E^{b,d} && \text{-}\\
&0 \leq p^{b,-}_h \leq \underline{p} \cdot E^{b,d} && \text{-}\\
&E^b_1 \leq E^b_{H+1} && \textit{Periodicity (optional)}\\
&\sum_{h=1}^{H} (p^{g,+}_h - (1 - \tau^{sh}) \cdot p^{ld,e}_h) \cdot \Delta h \leq 0 && \textit{Renewable share}
\end{aligned}$$

2.3 Numerical results

The objective of this section is to run parametric studies to examine the influence of the renewable share and the electricity tariff on the design and techno-economic indicators of the energy system.

2.3.1 Case study

The input PV production and demand time series are directly the 60 measured Ausgrid profiles introduced in section 1.1.4. Figure 2.2 shows the annual energy consumption variability (in MWh) between the 60 residential customers. In this case, the thermal energy demand approximately represents (in average) 20% of the total consumption for each customer.

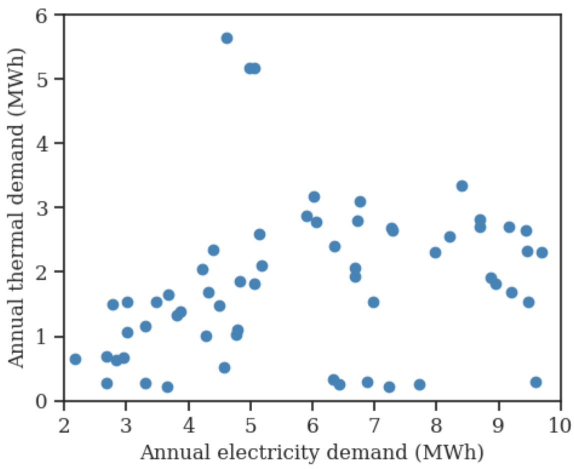


Figure 2.2: Annual energy consumption variability (in MWh) of the 60 Ausgrid profiles.

Storages	η^- [0-1]	η^+ [0-1]	η^{loss} [h^{-1}]	$\underline{e}$ [0-1]	$\overline{e}$ [0-1]	$\underline{p}$ [h^{-1}]	$\overline{p}$ [h^{-1}]	Lifetime [years]	Cost [€/kWh]
Li-ion	0.9	0.9	0.0005	0.2	0.8	1.5	1.5	12	300
TES	0.8	0.8	0.008	0	1	1.5	1.5	25	10
H2 tank	1	1	0	0	1	1.5	1.5	25	10

Table 2.1: Storage input technical and economical parameters

Input technical and economical parameters are given in table 2.1 and table 2.2, mainly based on [70] (using the authors "mode" values). The investment cost for solar panels is set to 1300 €/kWp. The "flate rate" price of electricity (taxes included) $c_h^{g,+}$ is fixed to 0.19 €/kWh following [71] for french residential customers in 2020. For the sake of simplicity, the feed-in tariff $c_h^{g,-}$ is set to zero in the rest of the study, meaning that the power grid is only seen as a generic dispatchable source.

Converters	$\eta^{e\rightarrow h}$ [0-1]	$\eta^{e\rightarrow h_2}$ [0-1]	$\eta^{h_2\rightarrow e}$ [0-1]	$\eta^{h_2\rightarrow h}$ [0-1]	Lifetime [years]	Cost [€/kW]
Heater	1	-	-	-	-	-
PEMFC	-	-	0.4	0.4	14	1700
PEME	0.3	0.5	-	-	15	1300

Table 2.2: Energy converters input technical and economical parameters.

The problem is modeled using Julia with the JuMP package [72], and solved with the IBM CPLEX 12.9 solver [59]. All the computations are run on a standard Intel(R) Core(TM) i5-7200U CPU @ 2.5GHz 2.7GHz computer.

2.3.2 Sensitivity of the renewable share

To assess the sensitivity of the renewable share over the results, the problem is solved for each of the 60 customers with a renewable ratio ranging from 0 to 1 (i.e. stand-alone DES: the energy demand is entirely supplied by solar panels). Again, remember that a 10 kW heater is assumed to be already installed, thus its size is not a decision variable of the problem.

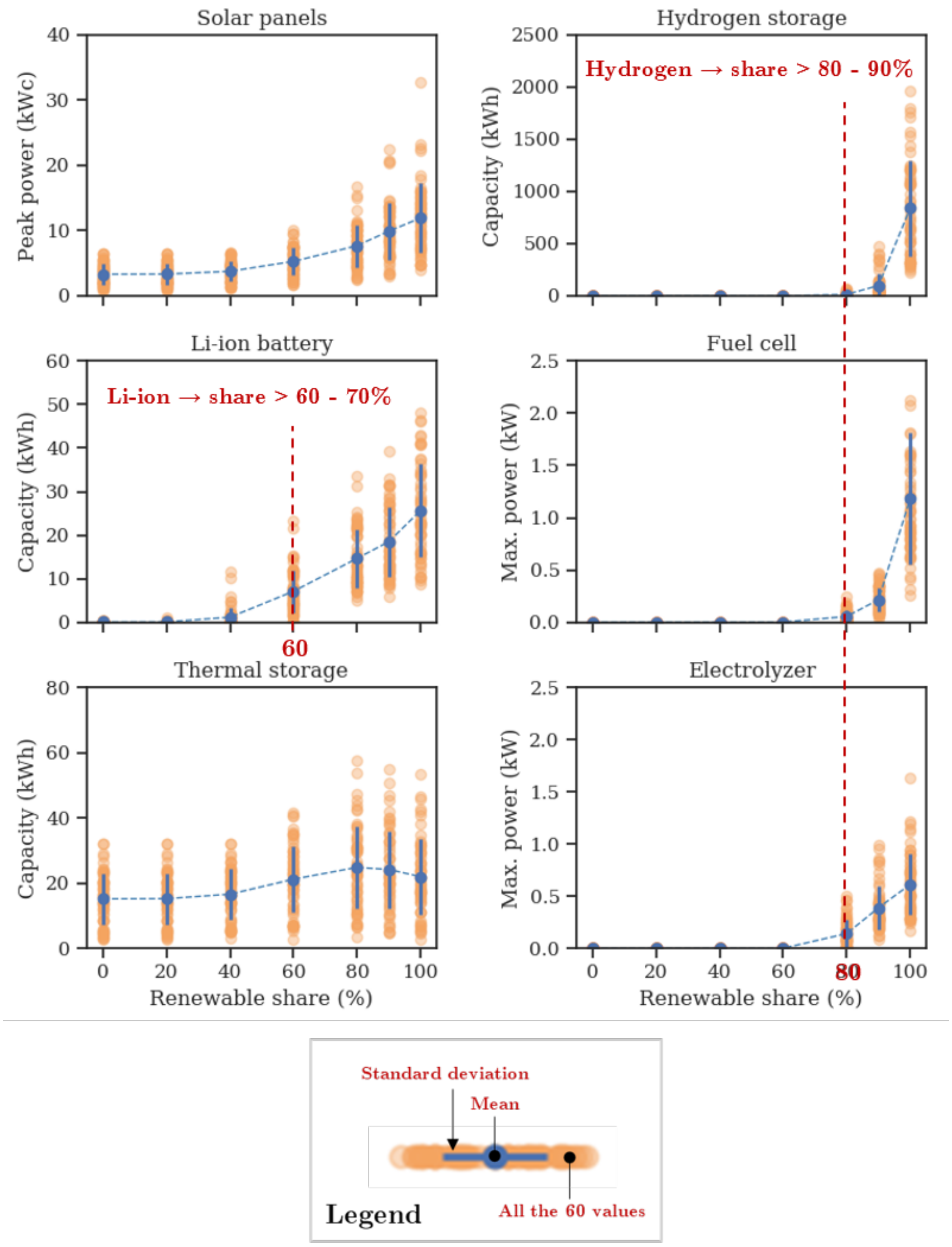


Figure 2.3: Sensitivity of the renewable share over the design.

Figure 2.3 shows the size of the assets as a function of the renewable share. All the 60 values are depicted in orange, while the error bar gives both the mean (blue marker) and the standard deviation (blue lines). The first observation is that only solar panels and TES are installed until the renewable share reaches 60%. Then, the Li-ion battery appears above this ratio while hydrogen units are only relevant when the share of solar production is higher than 80%. The equipment sizes increase exponentially along with the renewable ratio, except for the TES. Indeed, the cogenerated heat produced by the PEMFC and PEME is directly consumed, leading to lower thermal storage capacity to supply the thermal demand at the lowest cost. When the renewable share reaches 100% (i.e. the grid is no longer required), hydrogen storage provides seasonal flexibility while the battery and the thermal storage are used at shorter time scales (from hourly to monthly time scales for the TES). This latter statement is highlighted in figure 2.4 which depicts the Fast Fourier transform (FFT) of the storage state-of-charges (SoC) of a randomly chosen customer from the dataset. The x-axis is given in the log scale for the FFT. Note that the frequencies higher than the monthly time scale for both the Li-ion battery and the TES are artifacts of the model. Indeed, the spikes are due to the seasonal usage of the equipment (e.g., the TES is more likely to be used during winter than during summer), but no energy is stored over several months as shown in the "temporal evolution" graph. Readers interested in more elaborate analysis about flexibility issues could refer to [73].

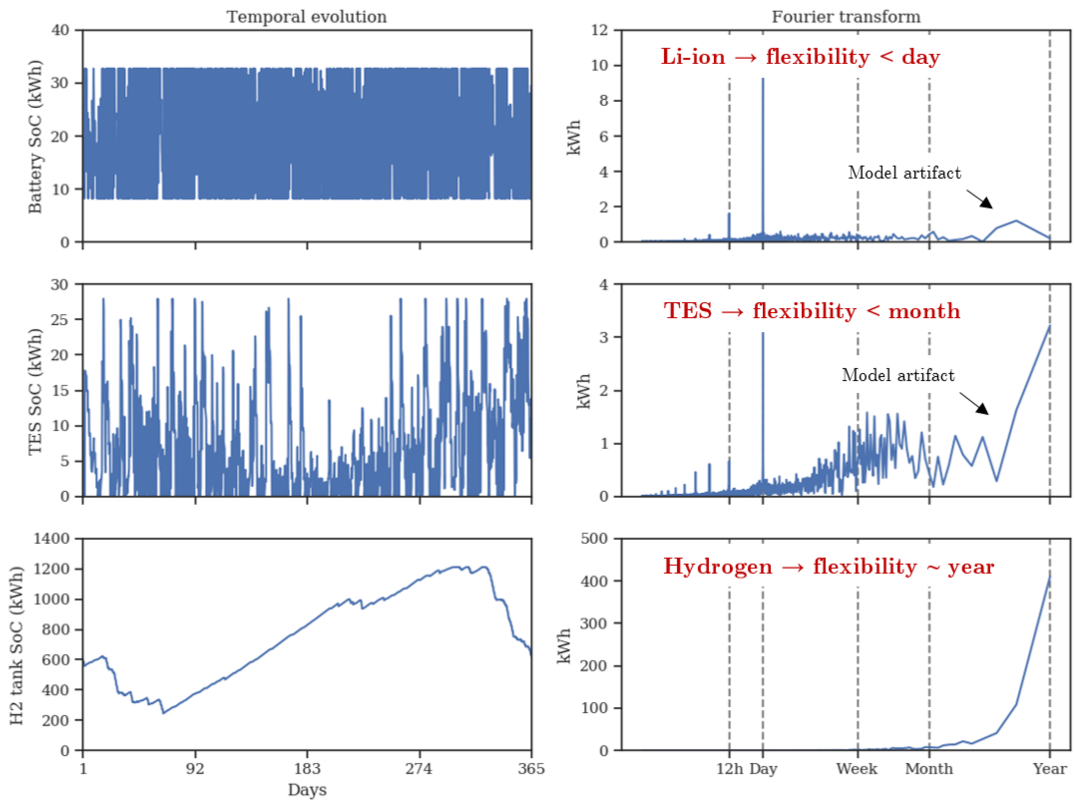


Figure 2.4: Fast Fourier transform (FFT) of the storage state-of-charges (SoC) for a randomly chosen customer from the dataset. The x-axis is given in the log scale for the FFT. Note that the frequencies higher than the monthly time scale for both the Li-ion battery and the TES are artifacts of the model. Indeed, the spikes are due to the seasonal usage of the equipment (e.g., the TES is more likely to be used during winter than during summer), but no energy is stored over several months as shown in the "temporal evolution" graph.

Obviously, increasing the renewable share comes at a price, depicted in figure 2.5. The latter displays as a bar chart, the mean annualized investment cost for each technology along with the operating cost from the grid. The mean and standard deviation of the total annual cost (i.e., involving both investment and operating costs) are also depicted with error bars in the figure. Note that the baseline cost (i.e., all the energy is supplied by the utility grid - "Ref" in the figure) is also displayed for comparison. As observed in the figure, reaching a 100% share of solar production is costly as the average total annual cost goes from 1.2 k€/y to 2.7 k€/y (+ 125%) and the integration of a low-cost dispatchable source may significantly alleviate the overall cost. For a standalone DES, the PV investment cost represents 38% of the overall cost while the Li-ion and hydrogen units equally account for 31%. Therefore, the cost of flexibility means to synchronize production and demands accounts for 62% of the total annual cost of the DES. In any case, the cost of the TES is negligible compared to the other costs. Note that the standard deviation is the highest for the standalone case. Finally, as shown in the figure, achieving 80% of renewable share (using a TES, a battery and PV) does not cost much more than the baseline case.

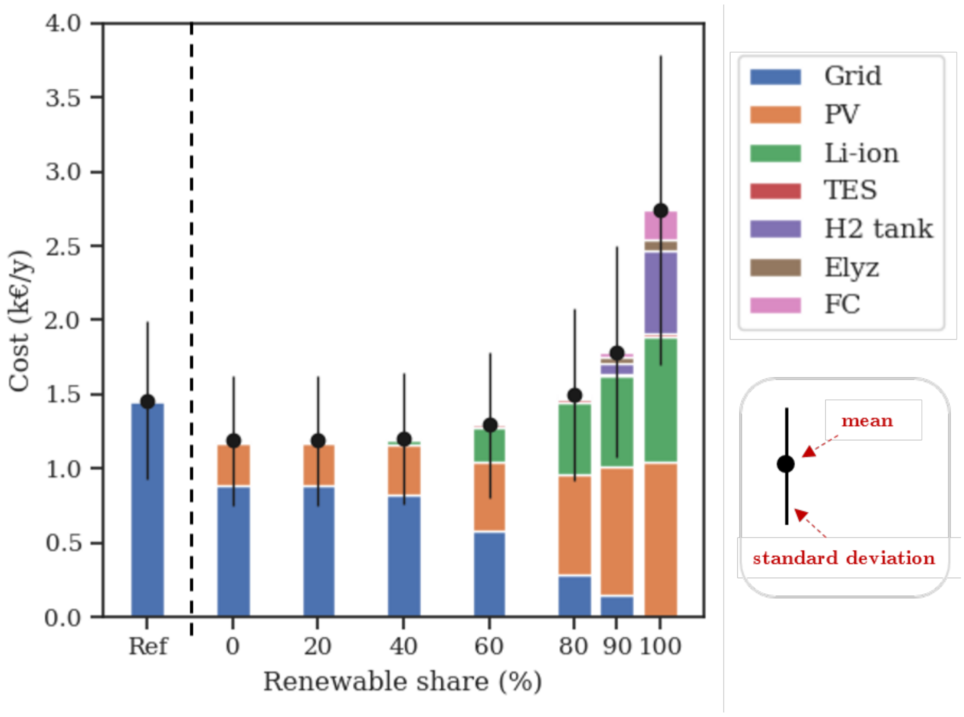


Figure 2.5: Sensitivity of the renewable share over the total annual cost. Thick bars correspond to the mean annualized investment costs for each technology along with the operating costs from the grid. The mean and standard deviation of the total annual cost are also depicted with error bars in black.

Next, the objective is to compare the multi-energy system to a battery-only solution (i.e., the toy-example including the thermal demand and the heater) to achieve the same requirements. Figure 2.6 shows the average annual cost as a function of the renewable share (only the error bars are plotted for clarity). The "without coge" case corresponds to the multi-energy system where the cogenerated heat from hydrogen is not recovered (i.e., PEME and PEMFC thermal efficiencies are set to zero). In this way, the additional value brought by cogeneration can be evaluated.

Results show that the mean cost value of the battery-only solution is significantly higher than the hybrid counterpart. This observation is even more relevant when high shares of solar production are needed: the mean cost difference goes up from 12% to 128% without the utility grid. For low values of renewable share, the only difference comes from the TES as no hydrogen assets are installed. Hydrogen cogeneration only makes sense when hydrogen units are actually installed - that is, with renewable share between 90% and 100%. In this case, the cogenerated heat helps decrease the total annual cost and the difference goes up to 11%. Keep in mind that in this case study, the thermal demand only represents 20% of the total consumption over one year.

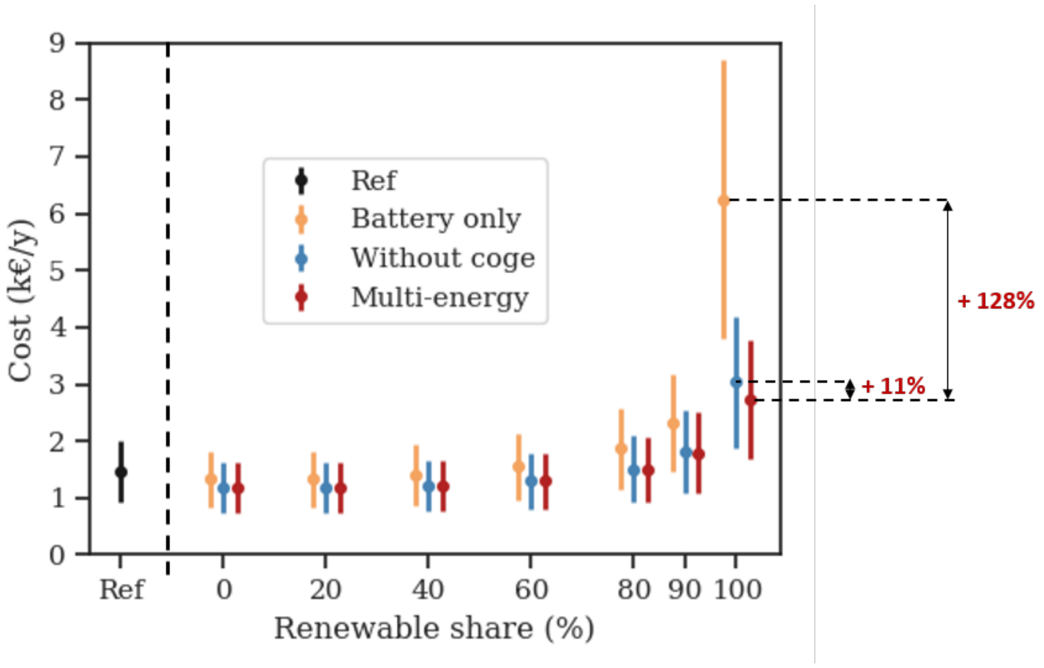


Figure 2.6: Comparison of the total annual cost between the multi-energy system, the battery-only solution and the multi-energy system where the cogenerated heat from hydrogen is not recovered (which corresponds to the ”Without coge” case in the figure).

2.3.3 Sensitivity of the electricity tariff

The objective of this section is to determine the cost of electricity needed to install Li-ion batteries and hydrogen systems in a profitable manner (from an economic point of view only). To this end, the renewable ratio is set to zero and the tariff of electricity is progressively increased from 0.19 €/kWh to 5 €/kWh. Figure 2.7 shows the sizing values as a function of the electricity tariff. As observed in the figure, the Li-ion battery and hydrogen units approximately emerge when the electricity tariff is greater than 0.3 €/kWh and 0.6 €/kWh (i.e., 3 times the current cost), respectively.

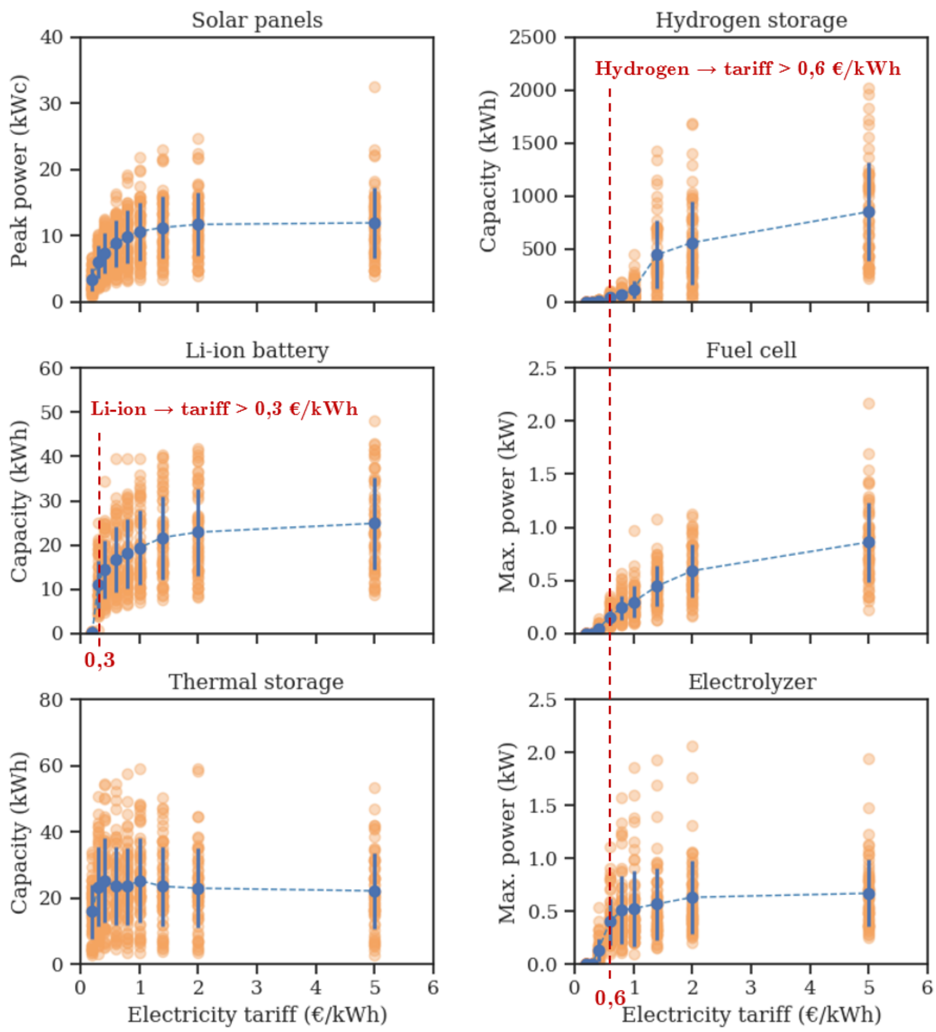


Figure 2.7: Sensitivity of the electricity tariff over the designs.

These values are then reported in figure 2.8 which shows the renewable share (computed *a posteriori*) as a function of the electricity tariff. Results show that the renewable share values are consistent with the results depicted in the previous section. In other words, to reach 70% and 90% of renewable share while being economically optimal, the electricity tariff should be higher than 0.3 €/kWh and 0.6 €/kWh, respectively.

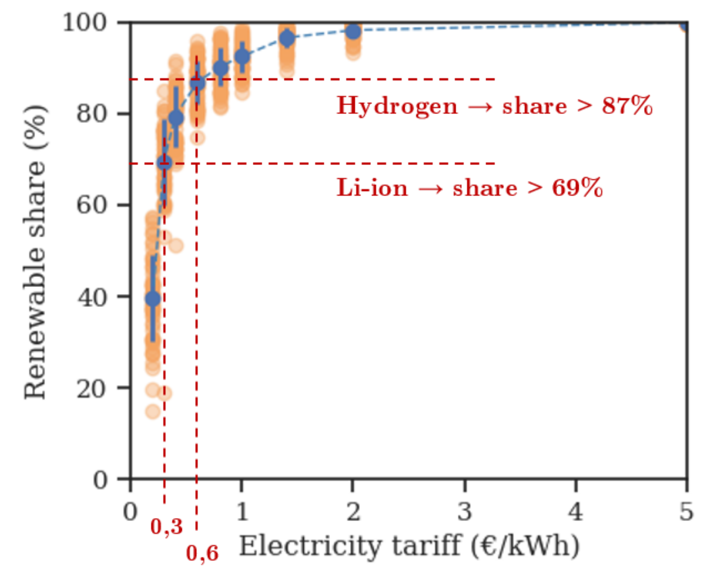


Figure 2.8: Sensitivity of the electricity tariff over the renewable share.

2.4 Discussion and conclusions

The value of multi-energy systems and seasonal storage to supply residential customers with a high share of solar production has been addressed in this chapter. First, the mathematical formulation of the problem was described in a deterministic framework. The objective of this part was threefold: i) such LP optimization problems are easily solved with standard solvers so that fast parametric studies can be run to identify global DES design and techno-economic trends; ii) the result of the deterministic problem with perfect foresight of the input parameters gives the best solution (but unreachable in reality) for a given scenario, which is a valuable tool to understand the topology of the problem; iii) the stochastic formulation introduced in chapter 4 is directly based on this formulation.

Then, the problem was solved over 60 real residential customer energy demands and production profiles to illustrate the relevance of the approach. The results show that under given assumptions, solar panels and TES are always installed, regardless of the value of the renewable share constraint. On the contrary, batteries only emerge whether the renewable share is greater than 60% or the electricity tariff exceeds 0.3 €/kWh, while hydrogen units only appear whether the renewable share is greater than 80% or the electricity tariff is higher than 0.6 €/kWh. In this latter case, hydrogen provides seasonal flexibility while the battery and the thermal storage are used at shorter time scales. Furthermore, the value of the multi-energy system over battery-only solution was clearly depicted in this case as the cost difference goes up to 128% when autonomy is required. Finally and without any surprise, the total annual cost increases exponentially with the renewable ratio and the integration of a low-cost dispatchable source may significantly alleviate the overall system cost.

Of course, quantitative results must be vigorously discussed and taken with caution as they highly depend on the case study, the input parameters and the modeling assumptions made in this work. Indeed, the results might be significantly different with more complex market designs (e.g., participation in local flexibility markets), or by allowing demand-side management strategies for instance. These latter aspects were out of the scope of this work. Also, another obvious limitation pertains to the deterministic nature of the model while the real problem is profoundly stochastic. Indeed, the future electricity tariff, the energy demand and production values are not perfectly known when sizing and operating the DES. Thus, the design values might be different to reach the same requirements, depending on the decision-maker's risk attitude. This latter limitation will be addressed in the second part of the thesis. Furthermore, even if the technological model granularity of the assets is consistent with the literature, they remain simple and their validity might be discussed. As explained in chapter 1, a good practice to overcome this limitation is to build a simulator with great technological details to evaluate the design results. Without this assessment phase, quantitative results must be considered with great caution.

Despite the aforementioned limitations, the mathematical formulation developed in this work seems to be a useful approach to identify global trends and evolution, concerning such multi-energy systems. This might be seen as the first important step before going into more realistic modeling to properly design the system according to a set of techno-economic requirements.

Part II

A stochastic framework for the design of multi-energy systems

Chapter 3

Generation of synthetic energy demand and PV production profiles based on Markov chains

Highlights

- Introduction of a simple method based on Markov chains to generate both long- and short-term synthetic scenarios of energy demand and production.
- The model recovers the main statistical features of the initial dataset while introducing temporal variability between scenarios.
- The model is not suitable to address epistemic uncertainties (i.e., only the power levels and daily patterns from the initial dataset are recovered).

Contents

3.1 Introduction

Since the deterministic model has been developed, the next step of this thesis is to include uncertainties (i.e., energy demand, renewable production, electricity tariffs) in both the design and operation decision processes. On one hand, the design of DES under uncertainty might be based on stochastic programming optimization techniques [74, 75] (further explained in the next chapter) where a large number of scenarios are required. On the other hand, once the size of the assets has been fixed, short-term probabilistic forecasts might be needed by real-time operation strategies to optimize the power flows between the equipment under uncertainty. For instance, look-ahead control strategies solve, at each time step, a multi-stage optimization problem, based on several probabilistic forecasts, each of them associated with a given probability (more details are given in [18] and [76]). In both cases, a large number of data over multi-time scales are essential to accurately solve the problems.

Having said that, decision-makers and modelers often lack appropriate data to run the models, especially in a stochastic context. In many real case studies, no historical data are available or the dataset is of poor quality, over short periods. Therefore, decision-makers might come up with inappropriate design decisions while modelers do not have enough data to assess the design and control approaches they are implementing. To overcome these difficulties, *scenario generation* methods have been widely implemented in the literature [77, 78]. This work mainly focuses on the generation of synthetic solar production and energy demands (i.e. electricity and heat) profiles at an hourly time step.

3.1.1 Literature review

While short-term forecasting is a relatively new topic (late 20th century) driven by efficient real-time operation needs, long-term forecasting for energy systems has been studied for a long time [77]. Indeed, the latter has been used for decades, to anticipate the energy demand growth in order to plan future energy production and transmission infrastructures. However, the recent and strong development of VRE has led to new long-term forecast requirements where short temporal granularity (i.e. at an hourly time step) is needed to cope with the short-time scale variability of the production [32]. Also, as noticed by Hong et al in [77] "*another important step in the recent history, is the transition from a deterministic to a probabilistic point of view*": instead of single values, the output of probabilistic forecasts are probability distributions of the uncertain parameters.

Recently, Mavromatidis et al [78] draw a great review of uncertainty characterization for the design of distributed energy system, which is of first interest for this work. A large number of methods are documented for both the generation of solar production and energy demand profiles, and the readers could refer to this article for an in-depth discussion about the different approaches. The objective of this part is to summarize the main conclusions and provide a clear insight into the direction of this study. Therefore, the first observation from their review is that the generation method depends on whether or not historical data are available. These approaches can be classified into *top-down* (i.e., historical data are available) and *bottom-up* categories, respectively. While obtaining solar production data is relatively straightforward [79], the availability of energy demand measurements is generally rarer.

In the top-down case, the most frequent and easiest generation method is the use of probability distribution functions (PDFs), derived from historical profiles for each hour. Then, a scenario is built by sampling from the PDFs. The drawback of such a method is that the uncertain parameters are treated as independent random variables between consecutive time steps, which might lead to unrealistic behavior where the autocorrelation and periodicity of the initial dataset are lost. To overcome this issue, more sophisticated and hybrid methods have been developed such as autoregressive models [80], Markov approaches [81], and machine learning based methods [82, 83, 84] to name just a few. The latter is probably the most popular approach for both the production and energy demands when large datasets are available [85].

On the other hand, when the case study lacks adequate energy demand measurements (e.g., newly built buildings), model-based methods are usually implemented to generate profiles. The most common approaches are probably the use of ready-made Building Performance Simulation (BPS) tools (e.g., energyPlus [86], TRNSYS [87]), but other model-based techniques are also implemented (e.g., resistance-capacitance (RC) models [88], a stochastic model where the input parameters are characterized based on interview information [89]). More elaborate methods are derived for large-scale districts where the previous approaches might not be appropriate (creating a model for each building of a district is quite laborious...) [90, 91]. In the bottom-up case, uncertainty is added to the input parameters of the simulation. The drawback of these methods is that a non-negligible amount of time is usually required to get familiar with BPS tools and collect all the numerous input parameters. Thus, energy modelers who are only seeking a fast generation method to test their design and operation algorithms might be discouraged by these approaches.

3.1.2 Main contributions

The main objective of this work is to provide a unique and straightforward method to generate a large number of probabilistic energy production and demand profiles when historical measurements are available. The energy modeler point of view is deliberately adopted in this work. The focus is more on creating a dataset to test different DES design and operation algorithms rather than the scenario generation accuracy. Nevertheless, the last section will show that the proposed method can capture the main statistical features and variations of real data despite the method's simplicity. Also, another important aspect is that the generation approach can be used simultaneously to generate long-term scenarios and short-term forecasts for operation purposes. Hence, the method is intended for modelers seeking a simple generation approach without spending too much time in this phase.

Therefore, the method implemented in this work is based on Markov chains over representative periods. The model only requires historical measurements of the uncertain time series in order to provide a wide range of contingencies. The methodology is introduced in section 3.2. Next, the performance of the approach is demonstrated on a residential case study from the Ausgrid dataset in section 3.3. Finally, the discussion and conclusion are drawn in section 3.4.

3.2 Methodology

The uncertain parameters (i.e. energy demands and solar production) are modeled as discrete random variables over a probability space (π, Ω). The following work aims at providing a method to build the discrete sample space Ω where a scenario is a sequence of all the random variable realizations over a given horizon H.

3.2.1 Building the Markov chains

Therefore, as said previously in the introduction, the generation method is based on Markov chains over representative periods. A Markov chain is a stochastic model where the main ingredients are the *states* and the *transition matrix*:

- *States* are observable realizations of the underlying random variables. The finite set of observed states is called the *state space*. In our case study, they are derived from the energy demand and production measurements.

- The *transition matrix* is a probability matrix where each cell is associated with the probability of going from one state to another.

The first step of the methodology (step 1 in figure 3.2) is to identify representative periods from the initial annual dataset to account for the different time scales variability. The Markov chains will be later computed over these periods. Therefore, each month of the year is gathered to avoid seasonality issues. Then, for each month, week and weekend days are divided into two classes as the energy demand pattern usually depends on the working activity. Finally, each day is segmented into 23 hourly periods to account for intraday variability. Thus, 23 x 2 x 12 Markov chains will be computed from the historical dataset. The classification of the representative periods is depicted in figure 3.1. Note that the representative periods are arbitrarily set based on both statistical explorations of the initial dataset and intuitions of the authors. Any other classes could be adopted depending on the nature of the random processes. This latter issue is definitely a weakness of this approach, which is further discussed in the last section.

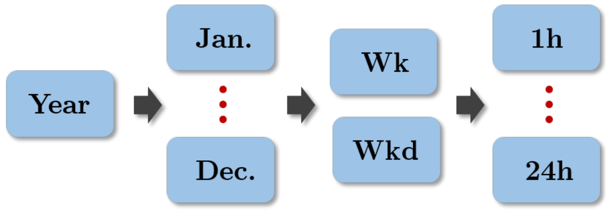


Figure 3.1: Representative periods classification to account for the different time scales variability.

Once the representative periods have been identified, a Markov chain is built for each hour where:

- *States* are aggregated (to keep the synchronicity) and normalized vectors of energy demands (i.e., electrical and thermal consumption) and production. To limit the number of state values for each hour, a given number k of relevant states is selected using the k-medoids clustering algorithm [92]. Therefore, each hour of each representative day is represented by $\{1, ..., k\}$ state values, associated with k vectors of aggregated and normalized power levels (step 2 in figure 3.2).

- *Probabilities* are computed based on the transition from one state to another between two consecutive hours. Thus, a 24 hours day is associated with 23 $k \times k$ transition matrices (one for each hour) with k state values for each hour (step 3 in figure 3.2).

3.2.2 Scenario generation

Based on the Markov model previously introduced, energy demand and production scenarios are generated by giving an initial state, a timestamp and the length of the horizon. In practice, the power values (which correspond to the Markov chain states) are sampled from the transition matrices where categorical distributions are built for each hour. The probability of each scenario π_s is given by the product of all the hourly probabilities. The generation process corresponds to the last step in figure 3.2.

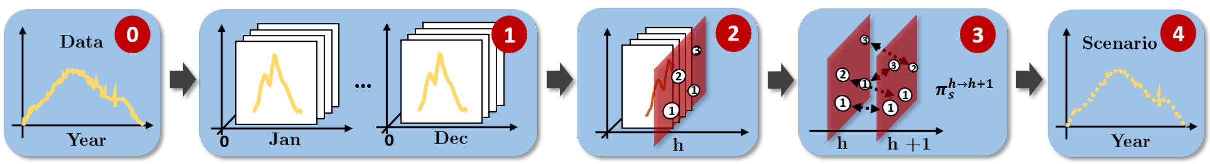


Figure 3.2: Description of the scenario generation method based on Markov chains: from historical data (0), days are classified into representatives week and week-end days for each month (1), for each hour, a given number of states is selected using the k-medoids algorithm (2), then the transition matrices based on the probabilities of going from one state to another between two consecutive hours are computed (3) and finally, synthetic scenarios are generated by giving an initial state, a timestamp and the length of the horizon (4).

3.3 Evaluation on a case study

The generation method is evaluated using the Ausgrid dataset where the 39th customer is arbitrarily chosen. Figure 3.3 shows the 3-year time series at an hourly time step for the electrical and thermal demands, in addition to the normalized solar production (in gray). Note that the first hour corresponds to the 1st of July as the season cycle is opposite to Europe.

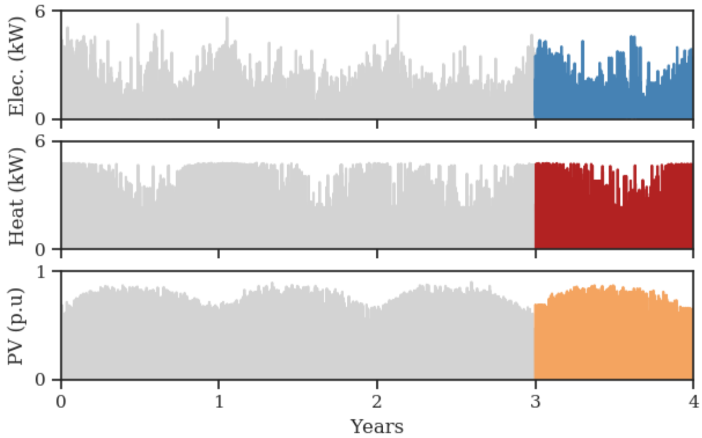


Figure 3.3: Overview of the 3-year time series from the 39th Ausgrid customer (in gray) and a one-year scenario generated with the Markov model (in color) for example.

While well-established metrics (e.g. root-mean-square error (RMSE), mean absolute error (MAE), etc.) are usually derived to assess the performance of short-term forecasting methods, the evaluation of long-term scenarios is less obvious at first glance. Therefore, following [81], [83] and [89] the evaluation for long-term scenarios will be based on a combination of both statistical and visual examination in comparison with the measured data.

3.3.1 Statistical assessment over the representative periods

To run the evaluation, Markov chains are built from the 3-year dataset of measured data. Then, 1000 scenarios of one year at an hourly time step are generated for the study. A single scenario is plotted in figure 3.3 with colored lines for comparison. A first general observation is that the shape of the profiles seems consistent with the measured data depicted in gray in the figure. This conclusion is also verified at a lower time scale as depicted in figure 3.4 and 3.5. Indeed, the latter show the comparison between the real data and the Markov model for both the week and weekend days of each month: the hourly mean values are depicted with a blue solid and red dash line for the model and the real data, respectively. All the values are also displayed in the figure backgrounds for both cases. As observed in the figures, it seems that the Markov model correctly reproduces the main statistical features of the initial dataset for each of the representative days. The seasonal issues are accurately addressed by the model as it follows the monthly variations of the real data. This latter observation is reinforced by comparing the power level amplitudes, in addition to the sunrise and sunset times of the different months. Note that for this case study, there are no major differences between the week and weekend days energy demand patterns. This latter observation might not be true with other residential customers.

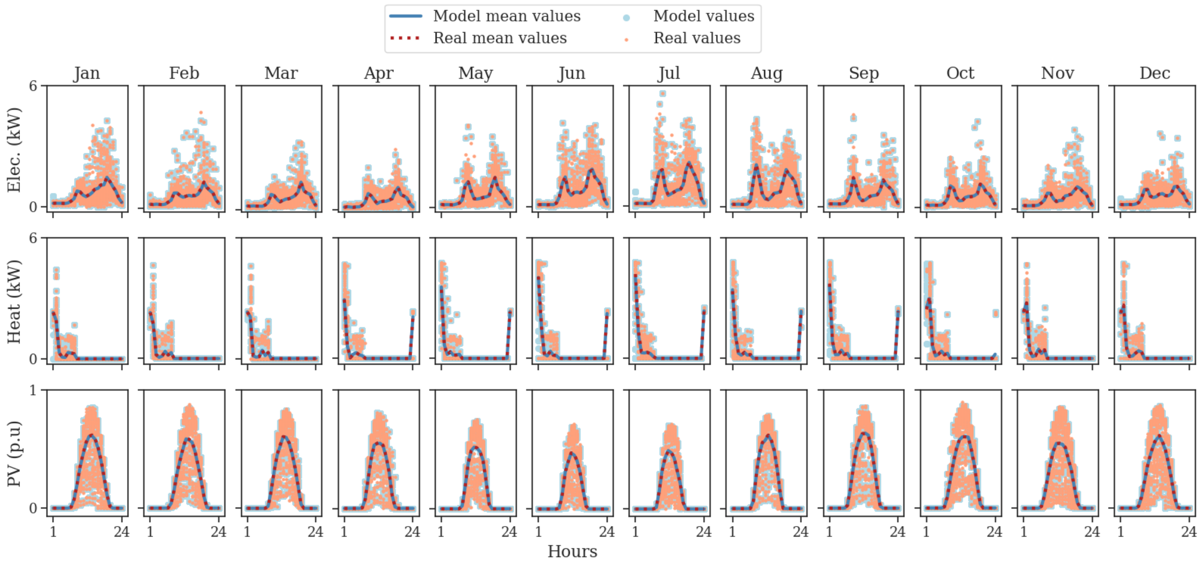


Figure 3.4: Comparison between the Markov model (in blue) and the real data (in red) for each **week** day of each month. Mean values are depicted with a solid and dash line for the model and the real data, respectively. All the values are given in the background of each figure for both cases.

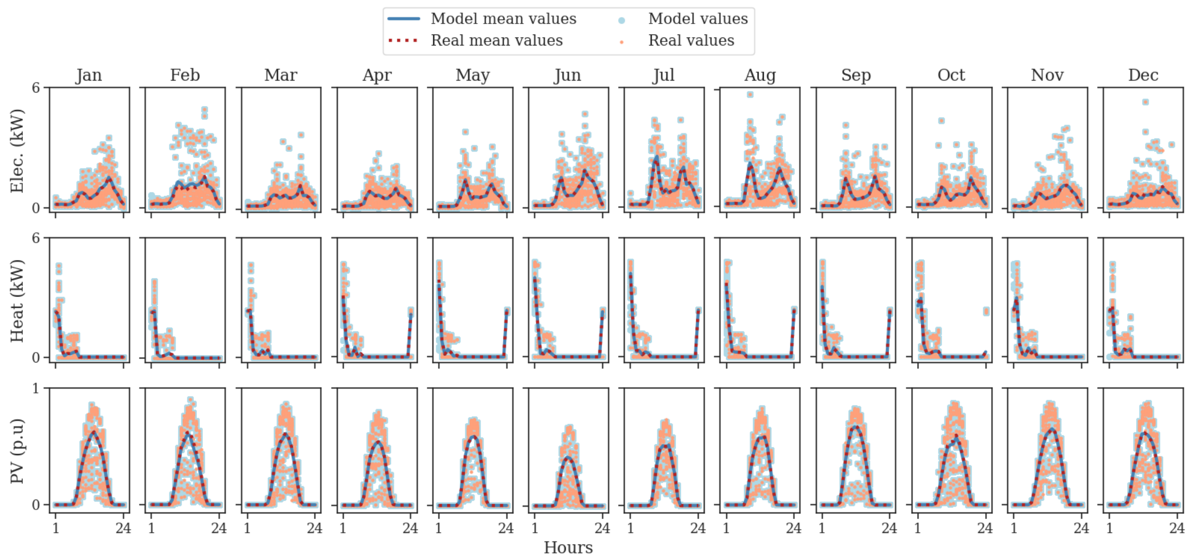


Figure 3.5: Comparison between the Markov model (in blue) and the real data (in red) for each **weekend** day of each month. Mean values are depicted with a solid and dash line for the model and the real data, respectively. All the values are given in the background of each figure for both cases.

3.3.2 Short time scale variability

Despite those statistical similarities, the Markov model still introduces short time scale variability from one scenario to another as shown in figure 3.6, where the energy demands and production are depicted over one week for 10 scenarios randomly chosen in July. Indeed, power values are not simultaneously the same between scenarios, which leads to a wide range of contingencies. This latter aspect is of first importance when dealing with the DES design and operation under uncertainties. Also, remember that each scenario is associated with a given probability which is computed thanks to the transition matrices (see section 3.2). Thus, the generation procedure is also suitable for short-term probabilistic forecasts, which can be later used by look-ahead control strategies to operate the DES (more details are given in chapter 4).

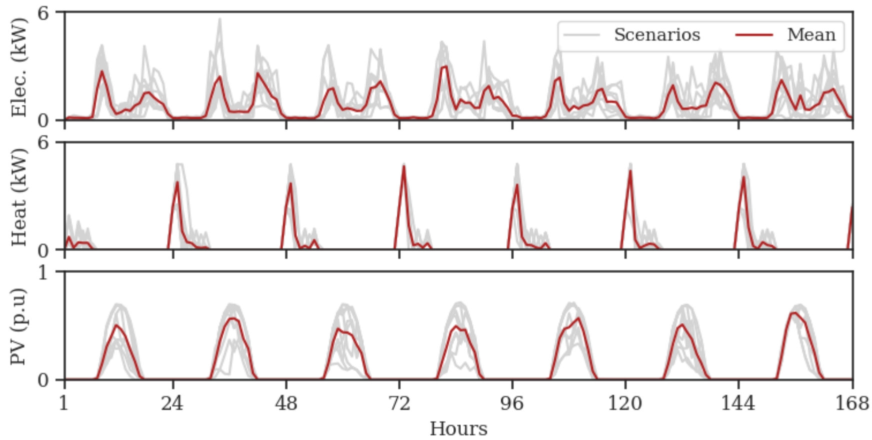


Figure 3.6: Short time scale variability over one week for 10 randomly chosen scenarios in July. The mean value is depicted in red.

3.3.3 Autocorrelation and duration curves

Autocorrelation refers to the correlation of a time series with a lagged copy of itself. The goal is to determine if the signal shows similarities between observations at different time lags. The result is given as a function of the delay (also called *lags* in figure 3.7). Despite the Markovian property attached to the generation method (i.e. the future state of the stochastic process only depends on the current state, without any memory of the past), the autocorrelation of the three variables is also recovered by the model as shown in figure 3.7. This might be explained as Markov chains are computed for each hour of representative days, leading to realistic power level sequences.

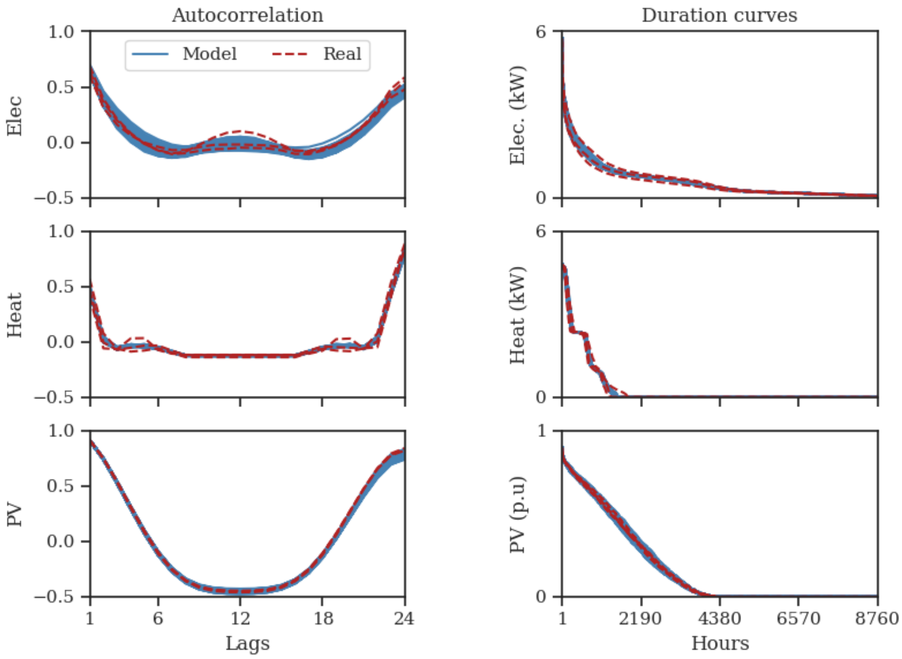


Figure 3.7: (Left) Autocorrelation of the three variables, and (Right) Load and production duration curves for both the synthetic scenarios (in blue) and the 3-year historical dataset (in red).

Finally, figure 3.7 also shows the *duration curves* of the three variables. With this representation, the values are sorted in descending order, which makes easier the comparison between the real data and the synthetic scenarios at a yearly time scale. The area under the curve corresponds to the total energy consumed (or produced) over the horizon. As shown in the figure, while model peak values are consistent with real data, the Markovian approach tends to generate scenarios with annual energy demands close to the average. Indeed, the model's blue curves are delimited by the real data. This latter observation is not verified for the production profiles. This is probably because the energy demand profiles have redundant patterns, leading to more peaked state probability distributions than for production. Indeed, despite the deterministic meteorological characteristics (i.e. the sunrise and sunset only depend on the position of the earth), the PV production is more likely to have random variations during the day, leading to more diversity in the generation process.

3.4 Discussion and conclusion

In order to generate synthetic scenarios for both long and short-term applications, a simple stochastic model based on Markov chains was presented in this chapter. First, the methodology was introduced where the Markov chains are computed over representative periods to account for the different time scales variability. Then, the method was applied to a residential case study where the objective was to build several energy demands and production scenarios. The results have shown that the main statistical features of the initial dataset have been recovered with this simple Markov model while introducing temporal variability to the annual time series. Finally, the last section has demonstrated that the Markovian approach is also suitable to generate short-term probabilistic forecasts, later used to control the DES.

The first limitation of this work comes from the classification procedure manually operated to identify the representative periods. Indeed, the performance of the Markov method is directly related to the expert knowledge concerning the structure and patterns of the initial dataset. Other approaches (mostly based on machine learning as in [83] for instance) do not require this first step and might be more relevant if little information is available about the stochastic processes. Moreover, although the Markov model introduces temporal variability into the scenario set, following the conclusion in [35], epistemic uncertainties are not addressed within this approach. Indeed, the model only recovers power levels and daily patterns that were already present in the historical dataset. When applied to a real case study, a strong assumption made by using this method is that the values of the uncertain parameters will remain the same in the future, regardless of their temporal variability. But what happens if the future energy demands increase or if the shape of the daily consumption changes? These latter issues are not properly addressed by only using the Markov model. This work aimed at developing a simple method to generate a large number of scenarios that will be later used to assess the different design and operation approaches. Decision-makers seeking quantitative and realistic results must spend a significant amount of time towards this generation phase.

Chapter 4

Design under uncertainty of a multi-energy system with seasonal storage: on the importance of the operation strategy in the design procedure

Highlights

- Introduction of a stochastic framework to study the design and operation of DES.
- Comparison of different design (i.e., mathematical programming and meta-heuristic approaches in a two-stage fashion) and operation (i.e., anticipative, rule-based, open-loop feedback control) strategies under uncertainties.
- Results show that the real-time operation strategy and the one embedded in the design procedure do not have to be strictly identical, as long as their levels of optimality are similar. This work helps quantify this notion.
- The mathematical programming approach (widely implemented in the literature) is relevant if, and only if, the DES is finally operated with a real-time policy that performs similarly to the anticipative operation strategy. This latter condition might be critical for large and complex energy systems.

Contents

4.1 Introduction

A common practice to assist decision-making under uncertainty is to run uncertainty and sensitivity analyses to assess the impact of stochasticity on the results [20, 55, 54]. In the former case, the objective is to observe the output variability of a model given multiple uncertain input realizations, while sensitivity analysis aims to identify the most influential input parameters regarding the output variability. This can be done using several well-known approaches as the Sobol or Morris methods [93, 94]. Following the modeling framework given in this work, both analyses can be performed either on the simulator or only on the design strategy (for example, references [20] and [55] only ran the analyses on the design optimization model). In the latter case, the authors built a deterministic optimization model on which the analyses were performed to evaluate the impact of uncertainty on the techno-economic indicators and the DES configuration. One of the objectives is to reduce the set of uncertain parameters in order to simplify the stochastic formulation. Despite the popularity of this approach, King and Wallace [75, 95] thoroughly explained why the results of such analyses must be used with caution. Indeed, the analyses are run on a deterministic model which does not reflect the real nature of the decision process (i.e., in reality, decisions are made without anticipating the future), so the sensitivity results are those of the deterministic model and their stochastic counterpart could be quite different. In any case, uncertainty and sensitivity analyses are made to provide insights into the impact and key drivers of uncertainty, but they do not deliver any design solution to hedge against them.

Having said that, this chapter deals with methods for optimization under uncertainty where the objective is to provide design values consistent with the risk aversion of the decision-maker. These techniques are far from new and two paradigms are mainly investigated in the literature: **stochastic programming** [96] and **robust optimization** [97]. The main difference between these two approaches is that the uncertain parameters are modeled as random variables over a discrete probability space with the former, while the latter uses uncertainty sets. Examples of such methods applied to the energy sector include [74, 98, 99, 100] for stochastic programming and [101, 102, 103, 104] for robust optimization.

In this work, the problem is formulated using the stochastic programming framework. As further explained in the following, the resulting stochastic design optimization problem is most of the time intractable, and simplifications are inevitable to come up with a solution. Among these simplifications, a common practice is to approximate the operation strategy (included in the design procedure), meaning that the DES will be controlled in real life with a policy different from the one used to design it. The main objective of this study is to challenge this latter hypothesis by comparing several design and operation methods under uncertainties.

4.1.1 Literature review

In the majority of the studies that apply stochastic programming to the design of DES, the optimization model is based on a single equivalent year where the uncertain parameters are modeled as random variables. To tackle long computation times, the information dynamic (i.e., the way uncertainties are revealed over time) is most of the time simplified, and uncertainties are assumed to unfold in only two stages: first, design decisions are made without perfectly knowing the future, then uncertainties are revealed, and *recourse* power flow decisions are made with perfect foresight over the operation horizon. The resulting model belongs to the well-known *two-stage recourse problem* category, widely discussed in the stochastic literature.

For instance, the authors in [74] and [98] implemented this technique for DES design, where a single large optimization problem is formulated and solved with standard commercial solvers. Other studies [99, 100] used decomposition methods for large-scale expansion planning problems where the number of variables is too large to be directly solved. Another widely used approach to solve the two-stage problem is to compute the sizing values thanks to a metaheuristic algorithm where the DES operation is simulated over a large set of scenarios and nested in the design loop [35, 105, 106]. As computational times are highly sensitive to the number of scenarios, *scenario reduction* methods are usually implemented to identify a reduced set of scenarios to statistically approximate the stochastic processes with a limited number of samples [100, 107, 108].

In all the previous works, **the operation strategy integrated into the design procedure cannot be implemented in real life**. Indeed, recourse decisions are made with perfect foresight over the entire operation horizon (i.e., one year), while realistic operation strategies cannot anticipate the future as they only have access to past and current observations. The operation might be based on forecasts, but in this case, forecasts are only computed from historical data, and decisions are made with limited foresight (i.e., usually one day). Therefore, the simplification of the operation strategy in the sizing procedure is a modeling approximation to facilitate the resolution of the design problem. However, this latter point has to be thoroughly discussed as the DES will be later operated in real life with a policy different from the one used to design it. In this case, how to be sure that the required performances will be met with a realistic operation strategy? Is the perfect foresight hypothesis (attached to the operation in the design method) appropriate for the design of DES? Should the operation strategies used for the design and in real life be the same? How sensitive are the design values to the operation policy? Despite their great importance for real-world applications, these design issues are, to the author's knowledge, rarely addressed in conjunction with realistic operation strategies. For example, the authors in [109] provided a brief comparison of different operation strategies with regards to the design but the study was carried out in a deterministic framework.

Another shortcoming of previous works concerns the evaluation of the stochastic solution. Indeed, the techno-economic indicators (e.g., total annual cost, renewable share) are usually computed over the same scenario set as the one used to design the DES [74]. To avoid any bias in the evaluation, the resulting design must be assessed on another set of scenarios, sharing the same statistical features but with different contingencies (e.g., synchronicity between the production and consumption, sequences of ”good” and ”bad” days). This process is called *out-of-sample assessment*, which is a common practice in the data science community, but rarely addressed in the energy planning field.

4.1.2 Main contributions

To address these shortcomings, the main objective of this work is to provide a comprehensive and joint comparison of different design and operation strategies under uncertainty for a multi-energy system with seasonal storage (depicted in figure 4.1). Especially, this study examines to what extent the aforementioned approximations (i.e., made for the purpose of the design), may lead to inaccurate results when operating the system in a realistic manner. To this end, the general framework introduced in chapter 1 is used: 1) first, the size of the assets is obtained by solving an integrated design approach where the approximated operation strategy Φ^d is embedded in a design loop to facilitate its resolution; 2) then, once the sizes have been fixed, the DES is evaluated on a common out-of-sample simulator with several realistic operation strategies Φ^a (i.e., which only have access to past and current information). The operation strategies Φ^d and Φ^a are not necessarily the same, so the objective of this work is to study the interplay between the design and the operation by varying the optimality level of the operation strategies in both phases (see figure 4.2 for a detailed representation of the methodology developed in this work). The only difference between the optimization and simulation phases is the way information is revealed over time: the design strategy may rely on the "perfect foresight" assumption, while the simulation is run with realistic operation strategies that only have access to past and current information. Note that the scenario sets Ω^d and Ω^a (i.e., time series of energy demands, generation and electricity tariffs) are different to avoid any bias in the evaluation. In practice, Ω^d could be historical measurements while Ω^a represents the future unknown values of the uncertain parameters.

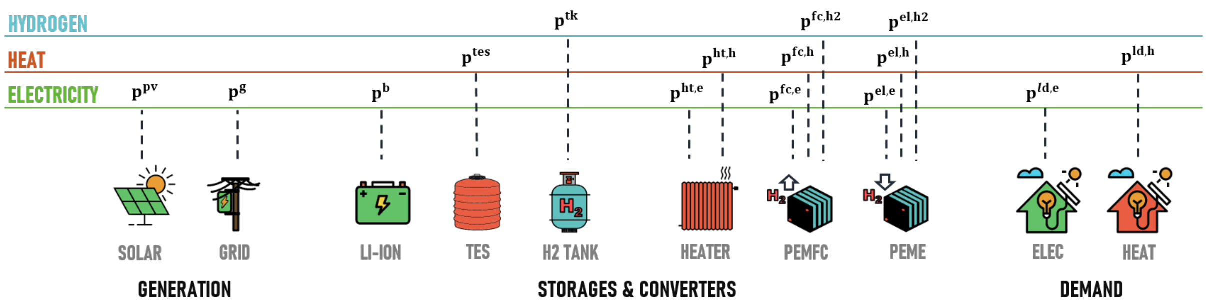


Figure 4.1: Schematic view of the multi-energy system with seasonal storage.

This work is organized as follows: section 4.2 shows the mathematical formulation of the design and operation problems where uncertainties are modeled as random variables. Then, section 4.3 describes the scenario generation method which is developed to build the sample space. Next, the resolution methods for both the design and operation of the DES are introduced in section 4.4. Then, the methods are applied and compared in section 4.5. Finally, section 4.6 provides in-depth discussions about the interaction between design and operation on a simplified case study, while the discussion and conclusions are drawn in section 4.7.

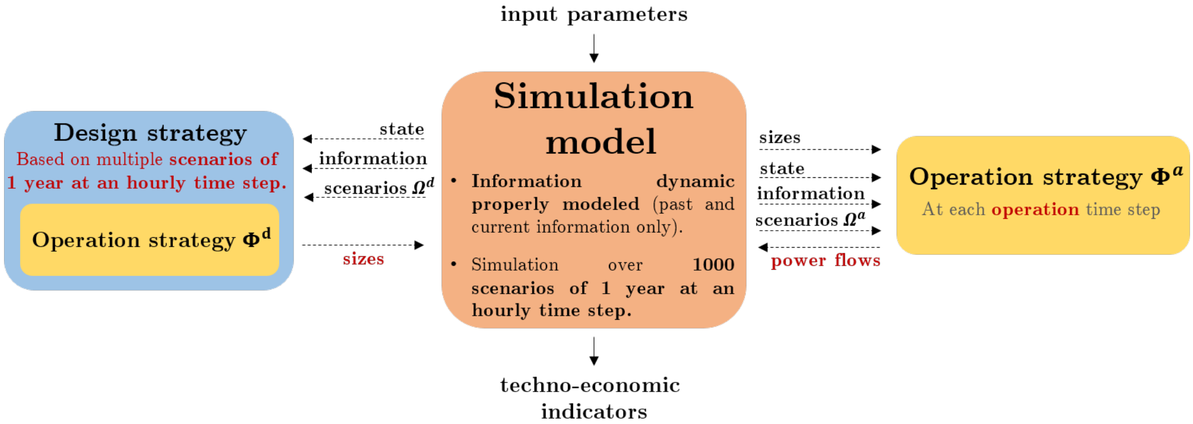


Figure 4.2: Detailed representation of the methodology developed in this work: 1) first, the size of the assets is obtained by solving an integrated design approach where the approximated operation strategy Φ^d is embedded in a design loop to facilitate its resolution; 2) then, once the sizes have been fixed, the DES is evaluated on a common out-of-sample simulator with several realistic operation strategies Φ^a (i.e., which only have access to past and current information). The operation strategies Φ^d and Φ^a are not necessarily the same, so the objective of this work is to study the interplay between the design and the operation by varying the optimality level of the operation strategies in both phases. Note that the scenario sets Ω^d and Ω^a (i.e., time series of energy demands, generation and electricity tariffs) are different to avoid any bias in the evaluation. In practice, Ω^d could be historical measurements while Ω^a represents the future unknown values of the uncertain parameters.

4.2 Mathematical formulation

In this section, the mathematical formulation of the stochastic problem is introduced. The model is mainly built on its deterministic counterpart introduced in chapter 2. The main difference is that operation variables become random variables as further explained in section 4.2.1. they are denoted with bold letters while their realizations are materialized by regular lower case letters. Only the main differences are introduced in this section, and the reader can refer to chapter 2 for more details about the energy system models.

4.2.1 Uncertainties

The electrical ($\mathbf{p}_h^{ld,e}$) and thermal ($\mathbf{p}_h^{ld,h}$) demands, the hourly solar capacity factor ($\mathbf{p}_h^{pv}$) and the electricity tariff ($\mathbf{c}_h^{g,+}$) are the uncertainties of the problem as their future values could not be predicted with perfect accuracy over the system lifetime. They are modelled as random variables over a probability space (π, Ω) and gathered in the vector $\mathbf{w}_h \in \mathbb{W}_h = \mathbb{R}^4$ (4.1). For the sake of simplicity, the electricity feed-in tariff ($\mathbf{c}_h^{g,-}$) is fixed to zero in the rest of the study.

$$\mathbf{w}_h = (\mathbf{p}_h^{ld,e} \quad \mathbf{p}_h^{ld,h} \quad \mathbf{p}_h^{pv} \quad \mathbf{c}_h^{g,+}) \tag{4.1}$$

In this work, a *scenario* $s \in \Omega$ is a sequence of 8760 realizations over one year, at an hourly time step $w_s = (w_{1,s}, \ldots, w_{H,s})$. The aim of section 4.3 is to build the discrete sample space Ω with the associated probabilities $\pi = (\pi_1, \ldots, \pi_S)$.

4.2.2 Decision variables

The decision variables for both the design and operation optimization problems are given as follows:

- The decision variables for the design are the size of the assets. They correspond to the maximum capacity of the storage systems and the maximum power of the energy converters, gathered in the vector $u^d \in \mathbb{U}^d = \mathbb{R}^6$ (4.2).

$$u^d = (E^{b,d} \quad E^{tes,d} \quad E^{tk,d} \quad p^{fc,d} \quad p^{el,d} \quad p^{pv,d}) \tag{4.2}$$

- The decision variables for the operation are the power flows controlled in the multi-energy system at every time step. They correspond to the charging and discharging power for the storage systems and the electrical powers for the energy converters, gathered in the vector $\mathbf{u}_h^o \in \mathbb{U}_h^o = \mathbb{R}^9$ (4.3). Because they depend on uncertainties, decision variables for the operation are also random variables.

$$\mathbf{u}_h^o = (\mathbf{p}_h^{b,+} \quad \mathbf{p}_h^{b,-} \quad \mathbf{p}_h^{tes,+} \quad \mathbf{p}_h^{tes,-} \quad \mathbf{p}_h^{tk,+} \quad \mathbf{p}_h^{tk,-} \quad \mathbf{p}_h^{fc,e} \quad \mathbf{p}_h^{el,e} \quad \mathbf{p}_h^{ht,e}) \tag{4.3}$$

4.2.3 Constraints

This section describes the constraints defining the admissible set of solutions for the decision variables previously introduced. The ”energy balances” and ”energy system model” constraints are not depicted in this section as they remain unchanged from chapter 2. The only difference is that operation variables become random variables. Special attention is paid to the flow of information (i.e., the way uncertainties are revealed over time) and the renewable share constraint. The entire stochastic formulation is given in a colored box for the toy-example at the end of this section.

Information structure

Particular attention is paid to the flow of information in the decision-making structure: in real-world problems, uncertainties unfold time step after time step and decisions are only made with past and current observations. As time passes, new information becomes available and the decision-maker could benefit from updated observations to make new decisions. A *stage* corresponds to a moment in time when the information is revealed, which is not necessarily related to the number of time periods. For instance, a ”two-stage” model only means that uncertainties are revealed in two stages which has nothing to do with the number of time steps in the problem.

In our case study, the first decisions made by the decision-maker before operating the DES, are the size of the assets u^d. These decisions are made without perfectly knowing the future uncertainty realizations over the equipment lifetimes. Next, when operating the system in real-time, values of the uncertain parameters are first observed and then, operation decisions are made. The same decision-making structure applies until the end of the system operation. The interplay between decision and uncertainties is depicted in (4.4).

$$u^d \rightsquigarrow \mathbf{w}_1 \rightsquigarrow \mathbf{u}_1^o \rightsquigarrow \mathbf{w}_2 \rightsquigarrow \mathbf{u}_2^o \rightsquigarrow \cdots \rightsquigarrow \mathbf{w}_H \rightsquigarrow \mathbf{u}_H^o \tag{4.4}$$

To formalize that operation decision variables only depend on past and current information, the so-called *non-anticipativity* constraint is introduced by equation (4.5) where σ is the σ-algebra generated by the input arguments (more details about the σ-algebra could be found in [19]).

$$\sigma(\mathbf{u}_h^o) \subset \sigma(\mathbf{w}_1, \dots, \mathbf{w}_h) \tag{4.5}$$

Share of solar production

In the following work, the renewable share constraint needs to be fulfilled according to a given risk measure. For example, if the risk measure is the expectation, this only means that the constraint has to be met in expectation but not in all cases. In this work, the so-called *conditional value at risk* (CVaR) is introduced as a coherent risk measure (for more information about the CVaR, see appendix A). The CVaR is convenient as it makes it possible to go from the expectation to the worst-case by only varying the β parameter from 0 to 1 (4.6).

$$\mathrm{CVaR}_\beta \left[\sum_{h=1}^{H} \Big(\mathbf{p}_h^{g,+} - (1 - \tau^{sh}) \cdot (\mathbf{p}_h^{ld,e} + \frac{\mathbf{p}_h^{ld,h}}{\eta^{e \to h}}) \Big) \cdot \Delta h \right] \leq 0 \tag{4.6}$$

4.2.4 Optimization problem statement

As in chapter 2, the integrated design objective is to determine the sizing and operation decisions in order to minimize the sum of both the annualized investment and operating expenditures.

Annualized investment cost

The annual capital cost depends on the equipment sizes which are the investment decisions of the design problem u^d, and the capital cost of each technology (4.7).

$$J^d(u^d) = \gamma^b \cdot c^b \cdot E^{b,d} + \gamma^{tes} \cdot c^{tes} \cdot E^{tes,d} + \gamma^{tk} \cdot c^{tk} \cdot E^{tk,d} + \gamma^{fc} \cdot c^{fc} \cdot p^{fc,d} + \\ \gamma^{el} \cdot c^{el} \cdot p^{el,d} + \gamma^{pv} \cdot c^{pv} \cdot p^{pv,d} \tag{4.7}$$

where the standard annuity factor γ is computed based on the expected lifetime of each technology (in years) with an interest rate of 4.5%.

Operating cost

At each time step, the operating cost is computed based on the energy exchanged with the external network (4.8).

$$J^o_h(u^d, \mathbf{u}^o_h, \mathbf{w}_h) = (\mathbf{c}^{g,+}_h \cdot \mathbf{p}^{g,+}_h - \mathbf{c}^{g,-}_h \cdot \mathbf{p}^{g,-}_h) \cdot \Delta h \tag{4.8}$$

where $\mathbf{c}^{g,+}$ is the tariff of electricity (€/kWh) and $\mathbf{c}^{g,-}$ the feed-in tariff (€/kWh), set to zero in the following work.

Problem statement

The resulting stochastic optimal design optimization problem is a nested problem where the operation is included in the design loop (4.9). The latter aims at finding the optimal operation decisions according to the decision-maker's risk attitude.

$$\min_{u^d} \; J^d(u^d) + \mathrm{CVaR}_\beta\Big[\widetilde{J}^o(u^d, \mathbf{w})\Big] \tag{4.9a}$$

$$\text{s.t.} \qquad u^d \in U^d \tag{4.9b}$$

where U^d is the set of admissible solutions for the design variables. The stochastic optimal operation problem is then given by (4.10).

$$\widetilde{J}^o(u^d, \mathbf{w}) = \min_{\mathbf{u}^o} \; \sum_{h=1}^{H} J^o_h(u^d, \mathbf{u}^o_h, \mathbf{w}_h) \tag{4.10a}$$

s.t.

$$\mathbf{x}_{h+1} = f(\mathbf{x}_h, u^d, \mathbf{u}^o_h, \mathbf{w}_h) \tag{4.10b}$$

$$\mathbf{u}^o_h \in U^o_h(u^d, \mathbf{x}_h, \mathbf{w}_h) \tag{4.10c}$$

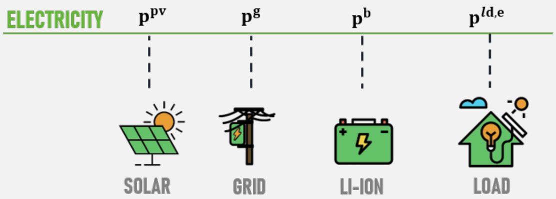
$$\sigma(\mathbf{u}^o_h) \subset \sigma(\mathbf{w}_1, \ldots, \mathbf{w}_h) \tag{4.10d}$$

where f is described by the state equations for storage systems. U_h^o is the set of admissible solutions for the operation which is related to the constraints introduced in section 4.2.3.

At this point, problem (4.9) is profoundly ”multi-stages” (and most of the time, intractable) as the information is only revealed time step after time step. Section 4.4 aims at introducing the two-stage approximation, in order to solve this problem.

Toy example - Two-stage stochastic formulation (deterministic equivalent)

The objective is to determine the size of the assets $u^d = (E^{b,d} \quad p^{pv,d})$ and the power flows controlled in the DES $u^o_{h,s} = (p^{b,+}_{h,s} \quad p^{b,-}_{h,s})$ at each time step $h \in \{1, \ldots, H\}$ and for each scenario $s \in \{1, \ldots, S\}$, in order to minimize the sum of both the annualized investment and operating expenditures. The uncertainties are given by $w_{h,s} = (p^{ld,e}_{h,s} \quad p^{ld,h}_{h,s} \quad p^{pv}_{h,s} \quad c^{g,+}_{h,s})$. More information about the linearization procedure are given in section 4.4.2.



The entire formulation is given as follows:

$$\min_{u^d, u^o} \; \zeta + \frac{1}{1-\beta} \cdot \sum_{s=1}^{S} \pi_s \cdot \alpha_s$$

subject to:

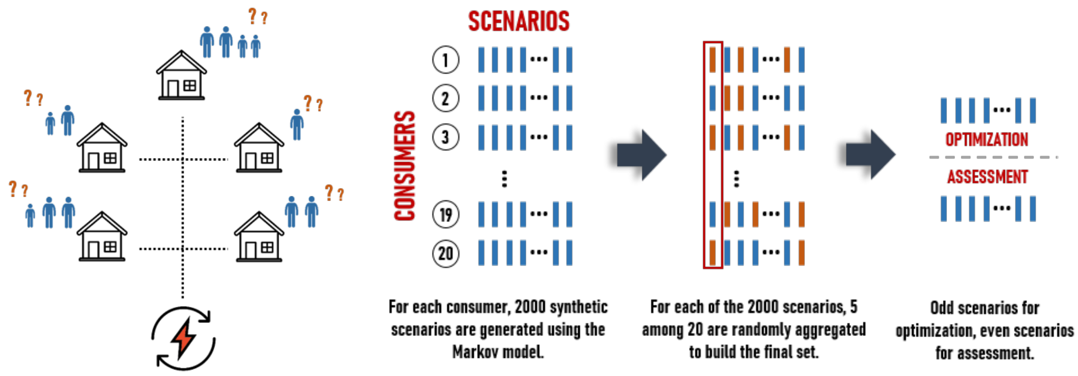
$$\begin{array}{ll}
\alpha_s \geq J^d(u^d) + J^o_{h,s}(u^d, u^o_{h,s}, w_{h,s}) - \zeta & \textit{Objective linearization} \\
\alpha_s \geq 0 & - \\
0 \leq E^{b,d} \leq \overline{E}^{b,d} & \textit{Size limitations} \\
0 \leq p^{pv,d} \leq \overline{p}^{pv,d} & - \\
p^{pv,d} \cdot p^{pv}_{h,s} + p^{b,+}_{h,s} + p^{g,+}_{h,s} = p^{ld,e}_{h,s} + p^{b,-}_{h,s} + p^{g,-}_{h,s} & \textit{Energy balance} \\
0 \leq p^{g,+}_{h,s} \leq \overline{g} & \textit{Power grid} \\
0 \leq p^{g,-}_{h,s} \leq \overline{g} & - \\
E^b_{h+1,s} = E^b_{h,s} \cdot (1 - \eta^{loss} \cdot \Delta h) + (\eta^{-} \cdot p^{b,-}_{h,s} - \dfrac{p^{b,+}_{h,s}}{\eta^{+}}) \cdot \Delta h & \textit{Storage model} \\
\underline{e} \cdot E^{b,d} \leq E^b_{h,s} \leq \overline{e} \cdot E^{b,m} & - \\
0 \leq p^{b,+}_{h,s} \leq \overline{p} \cdot E^{b,d} & - \\
0 \leq p^{b,-}_{h,s} \leq \underline{p} \cdot E^{b,d} & - \\
E^b_{1,s} \leq E^b_{H+1,s} & \textit{Periodicity (optional)} \\
\zeta' + \dfrac{1}{1-\beta'} \cdot \displaystyle\sum_{s=1}^{S} \pi_s \cdot \alpha'_s \leq 0 & \textit{Renewable share} \\
\alpha'_s \geq \displaystyle\sum_{h=1}^{H} (p^{g,+}_{h,s} - (1 - \tau^{sh}) \cdot p^{ld,e}_{h,s}) \cdot \Delta h - \zeta' & \textit{Constraint linearization} \\
\alpha'_s \geq 0 & -
\end{array}$$

4.3 Scenario generation

This section aims at building the discrete probability space (π, Ω) in order to accurately represent the uncertain parameters of the problem.

4.3.1 Case study

The distributed multi-energy system is assumed to be installed for a five residential houses district, newly built. We only know that the future inhabitants will behave as residential consumers but the exact number of people is *a priori* unknown (single person? a small family? a large family?). In order to estimate the future electrical and thermal energy demands, the scenarios generated in the next section are built upon a 3-year historical dataset of measured consumption for 20 residential consumers. This dataset comes from Ausgrid [10] and the authors in [11] identified a "clean dataset" where 20 consumers are taken. The total energy demands supplied by the multi-energy system is the aggregation of the five houses' consumption profiles (see figure 4.3).



(a) (b)

Figure 4.3: (a) The multi-energy system is installed for a five residential houses urban district newly built with unknown energy demands and production. (b) 2000 aggregated scenarios are built using the Markov model and split into two distinct parts: one set will be used for the design and the other for the assessment.

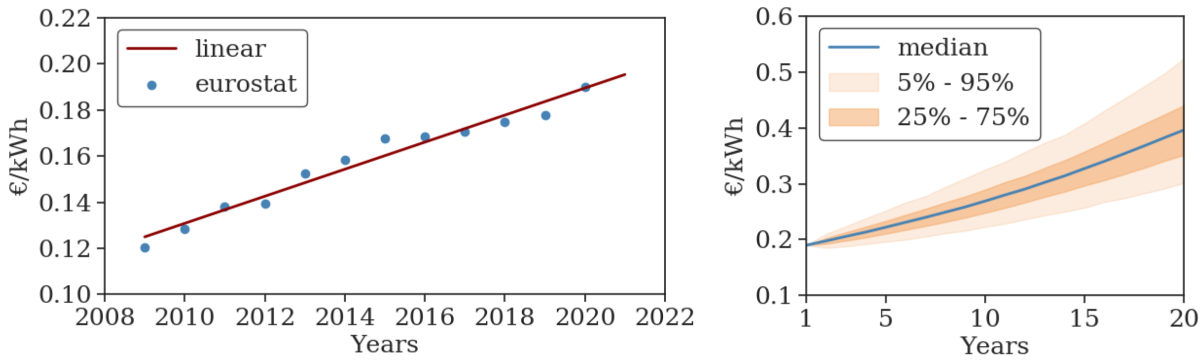
4.3.2 Generation process

Energy demands and production

Based on the Markov method developed in chapter 3, 2000 annual scenarios at an hourly time step are generated for each of the 20 consumers. Then, 5 among 20 are randomly aggregated to constitute the final set of energy demands and production scenarios used in this work (see figure 4.3).

Tariff of electricity

The price of electricity is another parameter that is usually concerned by uncertainty as it is unlikely to remain constant during the operational lifetime of the multi-energy system. As observed in figure 4.4 from data provided by Eurostat (European statistical office) [71], the "flat rate" price of electricity in France has increased from 0.12 €/kWh to 0.19 €/kWh (+58%) in only 12 years. From these observations, a simple statistical model is built which assumes a 4% increase per year with an additive white Gaussian noise of variance 4%. Then, 2000 scenarios projected over 20 years are randomly sampled from this model and the mean values over the 20 years for each projection are taken as inputs of this work similarly to [20] (see figure 4.4).



(a) (b)

Figure 4.4: (a) Evolution of the "flat rate" electricity tariff in France from Eurostat [71]. (b) Evolution of the electricity tariff over 20 years from the statistical model developed in this work. The mean values over the 20 years for each projection are taken as inputs of this study.

Scenario compilation

Energy demands and production profiles are randomly combined with electricity tariff scenarios, so a final set of 2000 scenarios with the same uniform probability is generated. Then, the final set is split into two parts: odd scenarios will be used for the design optimization phase (Ω^d) while even scenarios will be used for the out-of-sample assessment step (Ω^a). Figure 4.5 illustrates the annual energy demands (i.e., electricity and heat) diversity for both the design and the assessment sets.

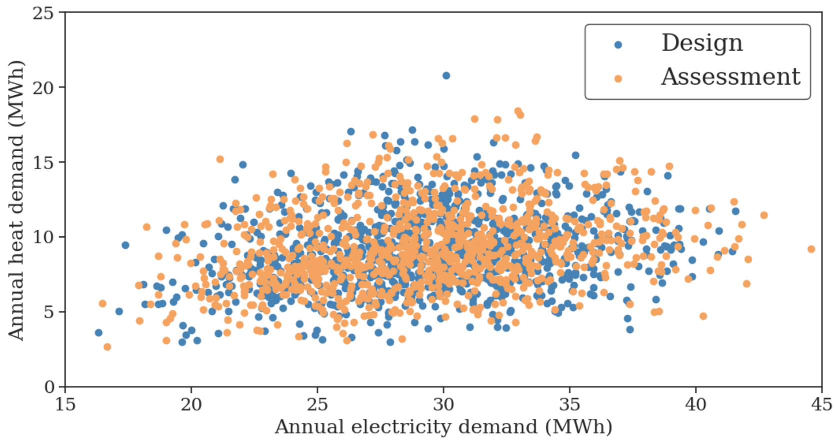


Figure 4.5: Annual energy demands variability for both the design and assessment sets.

The short time scale variability from one scenario to another is represented in figure 4.6 where the energy demands and production are depicted over one week for 10 scenarios randomly chosen.

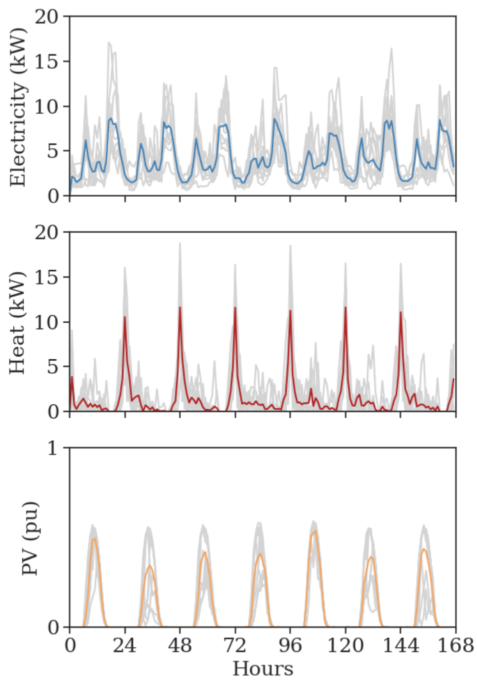


Figure 4.6: 10 scenarios of energy demands and production over one week generated using the Markov model (in gray). The colored lines are mean values.

4.4 Resolution methods

This section introduces the resolution methods for both the operation and design optimization problems, formulated in section 4.2. First, three operation strategies (i.e., rule-based, open-loop feedback control and anticipative) are introduced in section 4.4.1. Then, two integrated design methods (i.e., mathematical programming and metaheuristic) based on the "two-stage" approximation are presented in section 4.4.2. These approaches will be later compared in section 4.5.

4.4.1 Operation strategies

Unlike deterministic cases, the result of a stochastic optimal operation problem is a policy which is a sequence of mappings $\Phi = (\Phi_1, ..., \Phi_H)$ which gives at each time step the operation decisions as a function of the current state and available information $\Phi_h : \mathbb{X}_h \times \mathbb{W}_h \to \mathbb{U}_h^o$ [18, 19].

Two realistic policies are introduced in what follows: (i) a heuristic policy based on a set of rules, and (ii) a look-ahead method that solves a multistage optimization problem at each time step, based on a forecast of the uncertain parameters. Rule-based (RB) policies are widely implemented in real-life applications and in the literature because of their simplicity. However, when the system complexity increases (e.g., number of equipment, multi-energy carrier, complex market structure), it becomes more and more difficult to define a good set of rules to be close to the optimal operation. To address this shortcoming, look-ahead methods such as model predictive control (MPC) or its stochastic counterpart, open-loop feedback control (OLFC) [19] (also called *stochastic MPC*), are well suited for complex case studies. This comes at a price as an optimization problem needs to be solved at each time step, which might lead to a more complex control architecture. Note that many other optimized operation strategies could be found in the literature (e.g., stochastic dynamic programming, reinforcement learning) but their implementation is out of the scope of this work. The objective of this study is to compare a sub-optimal operation strategy (i.e. the RB policy described in the next section) with a smarter policy based on optimization (i.e. OLFC) to study the sensitivity of the operation strategy over the design. A smarter RB strategy could have been implemented but the objective of this work is to have two policies with different levels of optimality.

Rule based policy

A simple rule-based policy Φ^{rb} is implemented in this work: first, solar production is directly used to supply the electrical demand. Next, if there is a surplus of electrical energy, the charging priority is given to the Li-ion battery, followed by the H2 tank through the PEME, and the remainder is converted into heat through the heater. Otherwise, the battery is discharged firstly, then the PEMFC is turned on. The heat produced by the H2 cogeneration units is directly consumed to supply the thermal demand. If there is a surplus of thermal energy, the TES is charged. Otherwise, the TES is discharged and the heater is finally turned on if required.

Open-loop feedback control policy

As previously mentioned, OLFC is a well-known operation strategy which solves at each time step $h \in \{1, ..., H\}$ a multistage optimization problem, based on $s \in \{1, ..., S\}$ scenario forecasts of the uncertain parameters $\tilde{w}_s = (\tilde{w}_{h,s}, ..., \tilde{w}_{T,s})$ over a given horizon T (4.11).

$$\begin{aligned} &\Phi_h^{olfc} : \mathbb{X}_h \times \mathbb{W}_h \to \mathbb{U}_h \\ &\qquad (x_h, w_h) \to u_h \in \underset{u_h,\dots,u_T}{\arg\min}\, \mathrm{CVaR}_\beta \left[\sum_{t=h}^{h+T} J_t^o(u^d, u_t^o, \tilde{w}_{t,s}) + K(x_T) \right] \end{aligned} \tag{4.11}$$

The result is a sequence of optimal decisions $(u_h, ..., u_T)$ but only the first value is kept. The same procedure is repeated at each time step. The forecasts are generated with the Markov chain method introduced in section 4.3. For more details about OLFC, see [19]. A final cost $K(x_T)$ is typically added to the objective to avoid empty storages at the end of the look-ahead horizon.

The application of such a policy to seasonal storage is more challenging as the look-ahead horizon T is usually short enough to avoid long computation times (e.g., 1 - 3 days). Indeed, it would make no sense to repeatedly solve the problem over one year at each operation time step to deal with seasonal issues. In order to tackle this problem, a reference trajectory for the H2 tank state variable is added to the final cost and the deviation is penalized in the objective function (4.12). The reference trajectory comes from the two-stage design procedure where the deterministic equivalent problem over one year is solved. Then, the mean value of the deterministic H2 trajectories (one for each scenario over one year) is chosen as a reference target for the operation. The idea is similar to [110] and [111] but the target constraint (which represents the reference trajectory in their case) is somehow penalized in the objective function, leading to more flexibility in the operation.

$$K(x_T) = -k_1 \cdot E_T^b - k_2 \cdot E_T^{tes} - k_3 \cdot (E_T^{tk} - E_T^{tk,ref}) \tag{4.12}$$

where k_1, k_2 and k_3 are positive penalization coefficients calibrated by trial and errors (further explained in section 4.5). Without this final cost, it makes no sense for the controller to save energy for seasonal purposes and the tank would be used on a daily time scale. For the sake of simplicity, a simple linear penalization scheme is used but more complex penalization structures are possible to improve the performance of the operation strategy.

To summarize, OLFC is an optimized policy that has to be parametrized: the look-ahead horizon, the number of scenario forecasts, the penalization coefficients, the reference trajectory and the risk measure have to be fixed before operating the multi-energy system. Note that this chapter focuses more on design methodologies rather than optimal operation issues. Therefore, the parameters will be set manually by trial and error in section 4.5 but this issue could be a research question in itself for anyone interested in high-performance strategies.

Anticipative policy

The anticipative policy is also introduced as an unrealistic operation strategy. The latter finds the exact solution of the operation problem, assuming that all the information (i.e. realizations of the uncertain parameters) is available over the entire horizon for each scenario. The implementation of such "anticipative" (as the future is assumed to be perfectly known) policy is, of course, unfeasible in real life as uncertainties unfold progressively over time. However, its computation gives a lower bound to the operating cost (for each scenario) which is valuable information to measure the performance of any other realistic operation strategy.

4.4.2 Integrated design methods

As previously said in section 4.1 and 4.2, the resulting design problem is profoundly "multi-stages" and simplifications are inevitable to solve the problem. The most common simplification in the literature is the "two-stage" approximation where uncertainties are assumed to be revealed in only two stages. In this case, design decisions are made without knowing the future while the operation has perfect foresight of the uncertain parameters for each scenario. Therefore, two well-known algorithms (i.e., mathematical programming and metaheuristic) are introduced in this section to solve the two-stage optimization problem. They are depicted in figure 4.7 and further explained in the following.

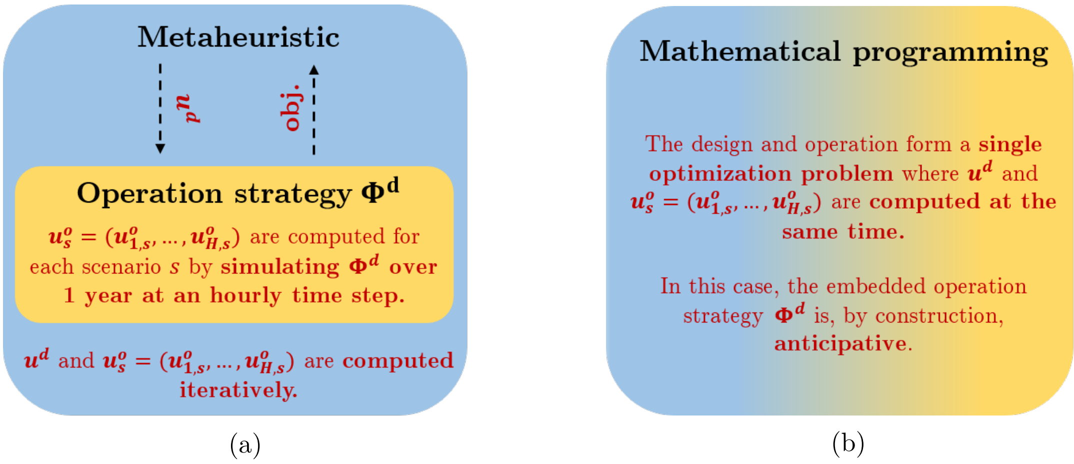


Figure 4.7: (a) With the **metaheuristic** approach, design and operation decisions are computed iteratively. The metaheuristic algorithm feeds a simulator with different system configurations (i.e., design decisions u^d) where the operation strategy is simulated over multiple scenarios of one year at an hourly time step. The process stops either when the objective has converged or the maximum number of iterations has been reached. (b) With the **mathematical programming** approach, the design and operation form a single optimization problem where both design and operation decisions are computed at the same time. Because of the two-stage approximation, the operation strategy embedded in the design procedure is indirectly anticipative.

Mathematical programming

The first approach is widely used in the literature [41, 74, 98, 99, 100], and implements a single large linear programming (LP) optimization problem where both the optimal design and operation decisions are computed at the same time. Therefore, the *deterministic equivalent* of problem (4.9) is formulated over a finite set of scenarios where the CVaR_β risk-measure is linearized according to [112]. As previously said in the introduction, this method leads to unrealistic operation policies as operation decisions are made with "perfect foresight" over the entire horizon. This simplification is a modeling approximation to facilitate the resolution of the design problem. The formulation is given by equation (4.13) where ζ and α_s are auxiliary variables introduced for the purpose of linearization.

$$\min_{u^d, u^o_{1:H,1:S}} \quad \zeta + \frac{1}{1-\beta} \cdot \sum_{s=1}^{S} \pi_s \cdot \alpha_s \tag{4.13a}$$

s.t.

$$x_{h+1,s} = f(x_{h,s}, u^d, u^o_{h,s}, w_{h,s}) \tag{4.13b}$$

$$u^d \in U^d, \;\; u^o_h, s \in U^o(u^d, x_{h,s}, w_{h,s}) \tag{4.13c}$$

$$\alpha_s \geq J^d(u^d) + J^o_{h,s}(u^d, u^o_{h,s}, w_{h,s}) - \zeta \tag{4.13d}$$

$$\alpha_s \geq 0 \tag{4.13e}$$

Unlike design variables, note that operation variables are indexed by the set of scenario s to take into account the non-anticipativity of the design in the "two-stage" structure. The same linearization technique is implemented for the renewable share constraint (see the toy-example for an illustration). The resulting problem is a single large LP optimization problem that can be solved with standard solvers (e.g., CPLEX [59]).

Metaheuristic

Another widely used method [24, 62, 113] is to use a metaheuristic algorithm to compute the design decisions: the DES is simulated over the set of design scenarios using one of the operation policies described above, and nested in a design loop. The size of the assets is computed through successive iterations. The algorithm stops either when the objective has converged towards a constant value or the maximum number of iterations has been reached. In this work, a niching genetic algorithm (also referred as *clearing*) [61] is implemented to compute the design decisions. The clearing is a niching elitist genetic algorithm that usually outperforms standard genetic algorithms on difficult problems with multiple nonlinear constraints and multimodal features. In this work, the renewable share and storage periodicity constraints are directly integrated into the objective function with penalty coefficients. The population size and the number of generations are respectively set to 50 and 100. Typical values for crossover and mutation rates are used (i.e., $pc = 1$ and $pm = 5\%$). In what follows, the algorithm is run multiple times in order to take the stochastic nature of the algorithm into account and to ensure the reproducibility of results.

The advantage of black-box algorithms is that no requirement on the model complexity (linear or not) is needed. Also, the embedded operation strategy might be the one used in real life, simulated over the set of design scenarios. However, the main drawback of this approach is that the number of iterations is usually very large to converge, with no guarantee of reaching optimality. Therefore, the operation strategy must be fast enough, and time-consuming policies like OLFC cannot be directly integrated into the design loop without any other simplifications (e.g., time step granularity, simulation horizon).

Scenario reduction

Solving previous methods using the 1000 design scenarios of Ω^d would result in intractable problems. Indeed, the total number of variables for the two-stage problem is approximately (without counting the auxiliary variables introduced for linearization):

$$\underbrace{6}_{\text{design decisions}} + (\underbrace{9 + 3}_{\text{operation decisions and states}}) \times \underbrace{8760}_{\text{hours}} \times \underbrace{1000}_{\text{scenarios}} = 105\,120\,006 \quad (4.14)$$

Therefore, *scenario reduction* methods are commonly implemented to tackle this concern [75]. The goal of such approaches is to identify a subset of scenarios $\Omega^r \subset \Omega^d$ to statistically approximate the initial set with a lower number of scenarios.

In this work, the reduction approach is based on clustering where the k-medoids algorithm is applied to a limited set of features (sum, maximum and first 4th statistical moments) previously extracted from each time series following [74] and [114]. The number of scenarios is fixed using stability tests [75]: the design methods are run with a growing number of scenarios until the objective function converges toward a constant value. In this work, 40 scenarios are needed, with each probability equal to the number of cluster assignments. Note that reduction methods is a fruitful research domain and more sophisticated techniques exist [100, 107, 108, 115], but their implementation is out of the scope of this work. More details about the scenario reduction technique implemented in this work are given in appendix B.

Therefore, the resulting problem has approximately 4 204 806 variables, which is computationally affordable for traditional solvers.

4.5 Application on the case study

This section is divided into four parts. In section 4.5.1, the perfect foresight hypothesis (attached to the mathematical programming design method) is challenged regarding realistic operation strategies. The goal is to discuss the validity domain of such an assumption. Next, section 4.5.2 introduces the metaheuristic method for poor performance policies with low computation time. Then, section 4.5.3 discusses the value of the stochastic solution compared to its deterministic counterpart. In the two previous sections, the objective is to entirely supply the consumption with the solar production (i.e., $\tau^{sh} = 1$) in a risk-neutral setting (i.e., $\beta = 0$ for both the objective and the renewable share constraint). Finally, a sensitivity analysis over the renewable share and the risk measure (i.e., β parameter) is carried out in section 4.5.4.

Storages	η^- [0-1]	η^+ [0-1]	η^{loss} [h^{-1}]	$\underline{e}$ [0-1]	$\overline{e}$ [0-1]	$\underline{p}$ [h^{-1}]	$\overline{p}$ [h^{-1}]	Lifetime [years]	Cost [€/kWh]
Li-ion	0.9	0.9	0.0005	0.2	0.8	1.5	1.5	12	300
TES	0.8	0.8	0.008	0	1	1.5	1.5	25	10
H2 tank	1	1	0	0	1	1.5	1.5	25	10

Table 4.1: Storage input technical and economic parameters

Input technical and economic parameters are unchanged from chapter 2. They are given in table 4.1 and table 4.2, mainly based on [70] (using the authors ”mode” values). The investment cost for solar panels is set to 1300 €/kWp.

Converters	$\eta^{e\to h}$ [0-1]	$\eta^{e\to h_2}$ [0-1]	$\eta^{h_2\to e}$ [0-1]	$\eta^{h_2\to h}$ [0-1]	Lifetime [years]	Cost [€/kW]
Heater	1	-	-	-	-	-
PEMFC	-	-	0.4	0.4	14	1700
PEME	0.3	0.5	-	-	15	1300

Table 4.2: Energy converters input technical and economic parameters

The problem is modeled using Julia with the JuMP package [72], and solved with the IBM CPLEX 12.9 solver [59]. All the computations are run on a standard Intel(R) Core(TM) i5-7200U CPU @ 2.5GHz 2.7GHz computer.

4.5.1 Evaluation of the mathematical programming approach with realistic operation strategies

In this section, the renewable share is fixed to 1 and the risk measure is the expectation. The first objective is to run the mathematical programming approach (also called "LP" in the following) to design the DES and then evaluate the solution with the realistic OLFC and rule-based operation strategies. Indeed, while most of studies solve the design problem assuming perfect foresight of the operation strategy, the main goal of this section is to examine to what extent this makes sense regarding realistic operation strategies which only have access to past and current information. In the following, Φ^a corresponds to the realistic policies while Φ^d is the anticipative strategy indirectly embedded in the design procedure to compute the size of the assets. Again, design values are computed using the reduced set of scenarios Ω^r, while they are evaluated with realistic policies over the 1000 assessment scenarios Ω^a to avoid any bias in the evaluation.

Therefore, the mathematical programming method is solved and the resulting design values are given in table 4.3. In the stand-alone case (i.e., $\tau^{sh} = 1$), the optimal cost option to reach autonomy is a combination of daily and seasonal storage along with multi-energy strategies to overcome the mismatch between solar production and energy demands. The total annual capital cost of the system is 14.7 k€ regarding the equipment sizes.

	Li-ion [kWh]	TES [kWh]	H2 tank [kWh]	PEMFC [kW]	PEME [kW]	PV [kWp]	Cost [k€/y]	CPU time [min]
LP	128	104	5085	3.5	3.8	68	14.7	133

Table 4.3: Design results with the mathematical programming approach. The risk measure is the expectation and the renewable share is equal to 1. The total annual capital cost and the CPU time (computational time required to solve the problem) are also added to the table.

More importantly are the results of the assessment phase run with both the optimized OLFC and RB strategies over the 1000 assessment scenarios Ω^a. The pending question is whether or not the renewable share constraint is met (on average) with realistic operation strategies despite the simplifications made for design purposes.

To this end, the OLFC strategy is parametrized by trial and error using the design scenario set Ω^d: once the sizes have been obtained, the OLFC parameters (i.e., the lookahead horizon T, the number of scenario forecasts S, and the penalization coefficients k_1, k_2 and k_3) are set manually by trial and error based on the expected value of the renewable share constraint computed on the design scenario set. They are given in table 4.4. Again, anyone seeking a high-performance policy needs to spend more time tuning these parameters. They can also be obtained by optimization based on a metaheuristic algorithm for instance. As previously said, the H2 reference trajectory comes from the design procedure and the risk measure is the expected value in this case.

Parameters	T [h]	S [-]	k_1 [€/kWh]	k_2 [€/kWh]	k_3 [€/kWh]
Φ^{olfc}	24	5	1e-3	1e-3	1e-2

Table 4.4: The OLFC parameters set manually by trial and error using the design scenario set Ω^d.

Figure 4.8 shows the statistical distribution of the out-of-sample renewable share as a histogram for both cases. As it can be observed, when the system is controlled with OLFC, most of the values (>700 occurrences over 1000) belongs to the upper interval (the renewable ratio is between 99.5% and 100%) and its expected value is close to 100% (see the red dash line in the figure which corresponds exactly to 99.6% in this case). In some rare events (<100 occurrences), its value is lower than 99% which is consistent with the expectation risk measure over the renewable share constraint. Note that, the operation policy could be even more improved to rigorously reach autonomy without shedding the loads. On the other hand, when the RB strategy is used, the distribution is much more flattened (values range from 94% to 100%) and the renewable share expected value is around 98%. Some would argue that this is quite a thin range and thus the optimization is relatively robust, but from a strictly mathematical point of view, the autonomy constraint is not met with respect to the project requirements.

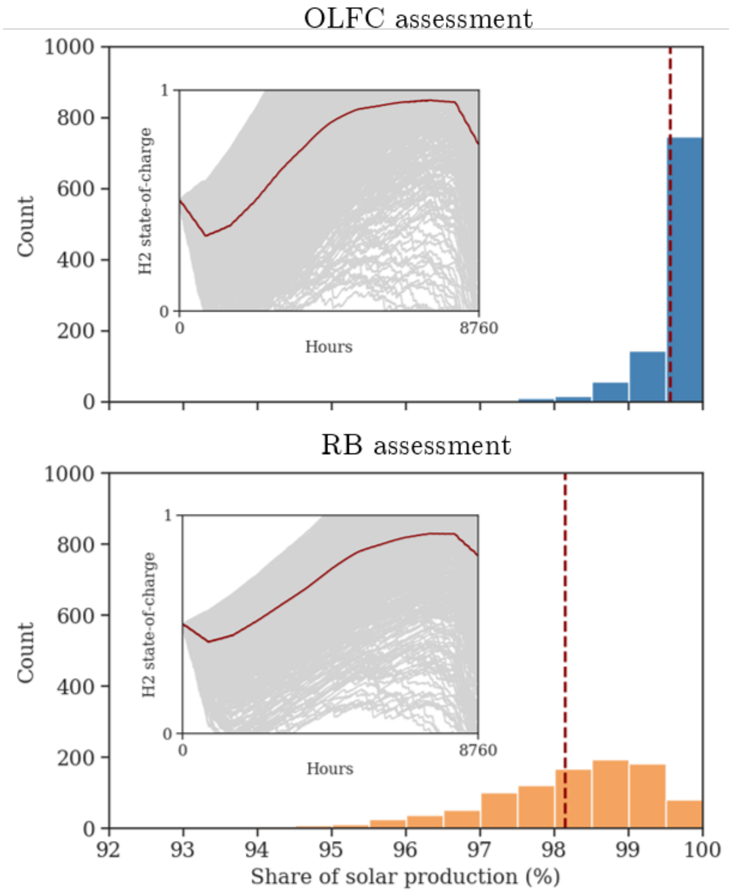


Figure 4.8: Histograms of the renewable share from the evaluation of the LP design results with the OLFC (top) and RB (bottom) policies. The dash red line is the expected value of the renewable share over the 1000 scenarios Ω^a, for both cases. The normalized H2 tank state of charge is also depicted for the 1000 scenarios in gray and its mean value is depicted in red.

This observation strengthens the fact that both operation strategies Φ^d and Φ^a must have similar performance (i.e. level of optimality). Otherwise, it can lead to underestimated (or overestimated) sizing values. As a result, **the mathematical programming method is only relevant if, and only if, the system is finally operated with a real-time policy that performs similarly to the anticipative operation strategy** (section 4.6 aims at quantifying this notion). Once the sizing is fixed, the only pragmatic way to verify that the constraints are met is to use out-of-sample simulations. This latter remark is even more important for complex DES (e.g., a high number of assets with complex market designs) because the implementation of a ”near-optimal” realistic operation strategy is difficult.

4.5.2 The metaheuristic approach for poor performance policies with low computation time

The remaining issue is how to provide relevant sizing values in the case where the decision-maker chooses to operate the DES with the sub-optimal RB policy. Indeed, OLFC is quite sophisticated for a real-world application and the decision-maker might choose the RB policy for its simplicity. In this case, the metaheuristic algorithm can be run with the nested RB strategy over the reduced set of design scenarios Ω^r. Note that the integration of the OLFC strategy into the metaheuristic approach would not have been possible without other simplifications (e.g, simplification of the temporal facet) for computational reasons (i.e., the simulation of the OLFC policy is too slow).

	Li-ion [kWh]	TES [kWh]	H2 tank [kWh]	PEMFC [kW]	PEME [kW]	PV [kWp]	Cost [k€/y]	CPU time [min]
LP	128	104	5085	3.5	3.8	68	14.7	133
Meta	149	585	1597	3.3	1.5	112	17.0	155

Table 4.5: Comparison of the design results with both the mathematical programming and metaheuristic approaches. The risk measure is the expectation and the renewable share is equal to 1. The total annual capital cost and the CPU time (computational time required to solve the problem) are also added to the table.

The design values are shown in table 4.5 in comparison with those of the mathematical programming approach. As shown in the table, the size of the PV (+65%), the Li-ion battery (+16%) and the TES (+463%) are increased while lower capacity hydrogen assets are installed. This difference might be explained as the RB policy is less effective at controlling the hydrogen systems to their full potential when coupling the energy carriers with each other. The comparison of the design results shows that the sizing solutions are completely different depending on the operation strategy embedded in the design procedure. Furthermore, the resulting design leads to a cost increase of 16% compared to the previous case. Hence, **the comparison between those two design methods allows assessing the potential cost reduction by using a near-optimal policy such as OLFC instead of sub-optimal strategies to reach the same requirements.** The out-of-sample renewable share distribution is depicted in figure 4.9. Unlike previously, the constraint is now entirely fulfilled with the RB policy and the expected value is 99.8%.

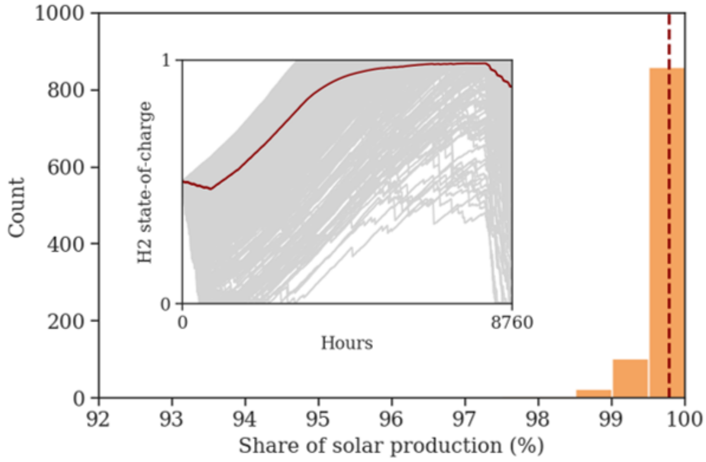


Figure 4.9: Histogram of the renewable share from the evaluation of the metaheuristic design results with the RB policy. The dash red line is the expected value of the renewable share ratio over the 1000 scenarios. The normalized H2 tank state of charge is also depicted for the 1000 scenarios in gray and its mean value is depicted in red.

As a conclusion from the two previous sections, the results could be summed up as follows:

- Most of the planning studies (in the literature) solve a single mathematical programming problem, but the only condition to make it a relevant design approach is that the DES is finally operated with a real-time policy that performs similarly to the anticipative operation strategy. Otherwise, the constraints might not be met.

- If the operation strategy is fast enough to be nested in the metaheuristic algorithm, this latter approach is preferred to avoid inappropriate design values.

- Both methods provide design values consistent with a given risk measure. The comparison between those two approaches allows assessing the gain of using optimized policies instead of sub-optimal strategies to control the DES.

- Whatever the design and operation method, this study highlights the importance of both the design and assessment phases to ensure that the techno-economic requirements are met with realistic operation strategies. The interaction between design and operation is again strengthened by this work.

Concerning computation performances, 2h15 is needed to solve the LP design problem which has approximately 4 204 806 variables, and 2h35 with the metaheuristic algorithm for a single run.

4.5.3 Limitations of deterministic designs

Since uncertainties are usually ignored in most studies, this section examines the value of the stochastic solution in comparison with its deterministic counterpart. The question is to what extent this approximation could lead to inaccurate results. In what follows, deterministic problems are solved using three different scenarios and evaluated on the out-of-sample simulator with the OLFC strategy.

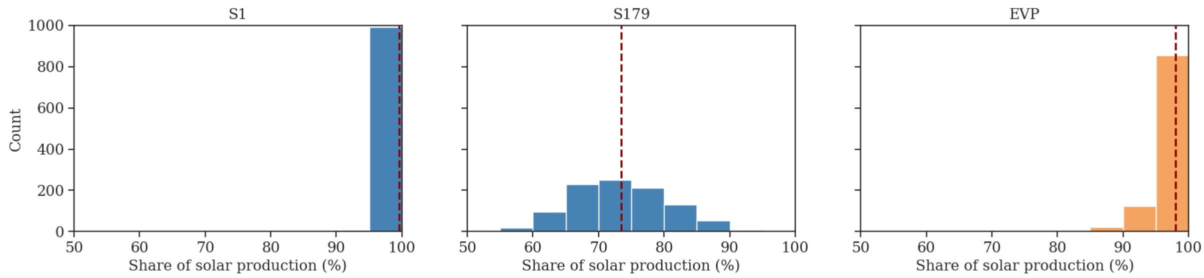


Figure 4.10: Renewable share statistical distribution for 3 deterministic cases: ”S1” and ”S179” correspond to the 1st and 179th scenario, while ”EVP” is the result of the *expected value problem* where the uncertain parameters are replaced by their expected value. The dash red line is the expected value of the renewable share ratio over the 1000 out-of-sample scenarios.

Figure 4.10 shows the renewable share statistical distributions and the sizing values are given in table 4.6: ”S1” and ”S179” are the 1st and 179th scenarios randomly selected for optimization, while ”EVP” is the result of the *expected value problem* further detailed in what follows. As observed, the resulting out-of-sample share of solar production distribution highly depends on the scenario selected for optimization: in the ”S1” case, the results are really close to those obtained with the stochastic approach, while the requirements are, by far, not satisfied with the 179th scenario. In this latter case, the sizing values are largely underestimated, leading to a renewable share expected value of around 75% and a total cost equal to 6.3 k€/years (-57% compared to the stochastic result). Thus, deterministic results could lead to serious inaccurate solutions depending on the scenario used for optimization. One could say that the scenario must be cleverly chosen according to the goal of the study. But what does ”good”, ”bad” or ”middle” scenario means in a design perspective for complex multi-energy systems? Also, how to select the adequate scenario to finally ended up with out-of-sample results consistent with the decision-maker risk aversion? It feels like these questions are not properly addressed with a deterministic point of view.

	Li-ion [kWh]	TES [kWh]	H2 tank [kWh]	PEMFC [kW]	PEME [kW]	PV [kWp]	Cost [k€/y]	CPU time [s]
S1	151	123	4700	3.0	2.1	65	14.7	40
S179	54	55	2027	1.8	1.9	30	6.3	40
EVP	117	99	3333	2.4	1.5	59	11.9	40

Table 4.6: Sizing values from deterministic methods with a renewable share equal to 1 for several scenarios.

Another common temptation is to replace all the random variables in equation (4.9) with their expected values, leading to a deterministic formulation with a single scenario.

This latter problem is known as the *expected value problem* (EVP) in the stochastic programming community [96]. In this case, the results are closer to the stochastic solution as depicted in figure 4.10 but the autonomy constraint is still not met: its expected value is 98% with some rare values reaching 85%. Consequently, despite shorter computational times (40 seconds against approximately 2.5 hours for the stochastic resolution), results from deterministic models might be hazardous and should be carefully treated as they could lead to underestimation of the total system cost with overestimated performances.

4.5.4 Renewable share and risk measure sensitivity analysis

Since the previous results have been obtained in a risk-neutral setting with autonomy requirement, the aim of this section is to vary these two parameters in order to study their consequences over the sizing values and total system cost: for the risk measure, only the expectation (i.e., $\beta = 0$) and the worst case (i.e., $\beta = 1$) are investigated in this work but any other analyses could be run by varying the β parameter.

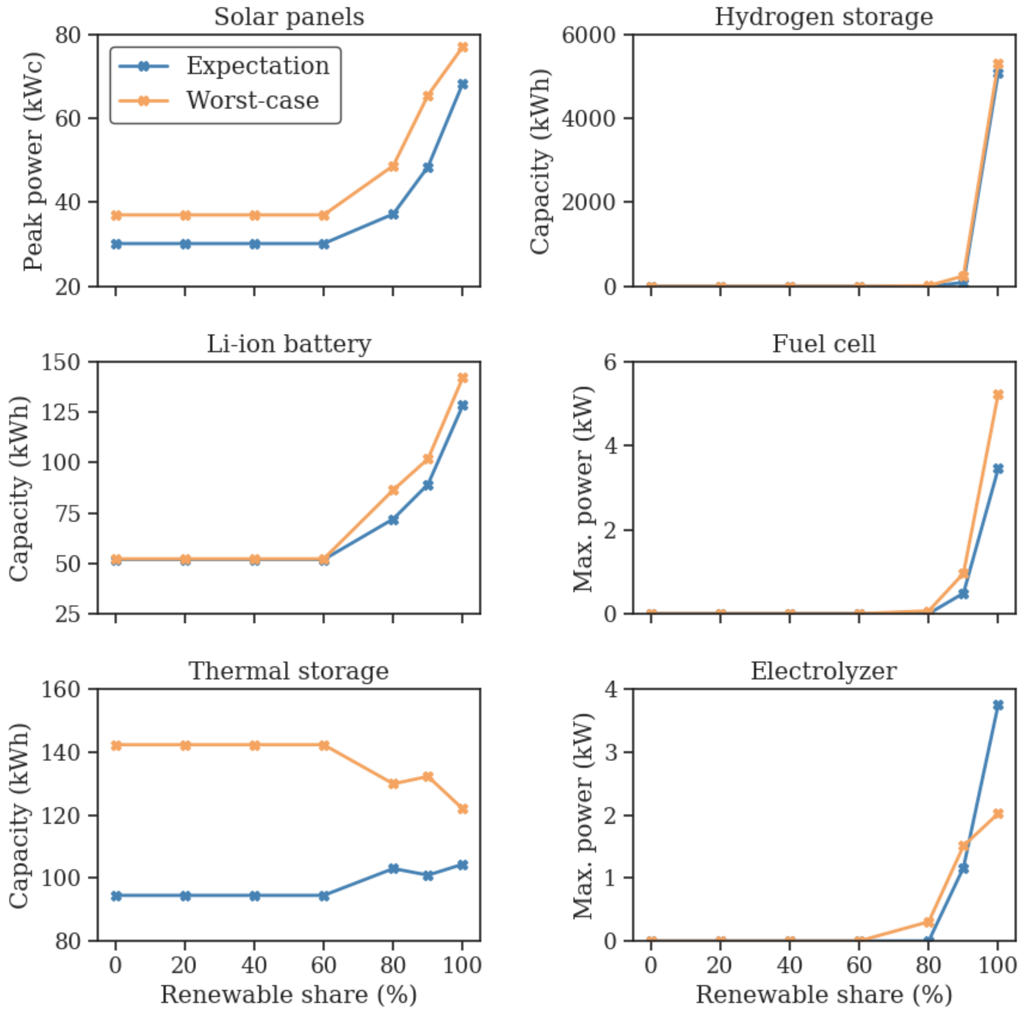


Figure 4.11: Sizes of the equipment for different share of solar production in the risk neutral and worst cases.

Figure 4.11 depicts the design values as a function of the renewable share for both risk measures. Unlike the deterministic results of chapter 2, a Li-ion battery is installed even with low renewable share constraints. This is probably due to the uncertainty attached to the cost of electricity over the system's lifetime. In the expectation case, a combination of 30 kWp of solar panels, 95 kWh of thermal storage and 52 kWh of Li-ion batteries is profitable up to 60% of renewable share, regarding the assumption of this work. Then, hydrogen only emerges since the renewable share is greater than 90%, which is consistent with the conclusion of chapter 2. Increasing the share of solar production comes at a price as the total cost exponentially increases between 60% and 100% and its value jumps from 7.8 k€/year to 14.7 k€/ (see figure 4.12).

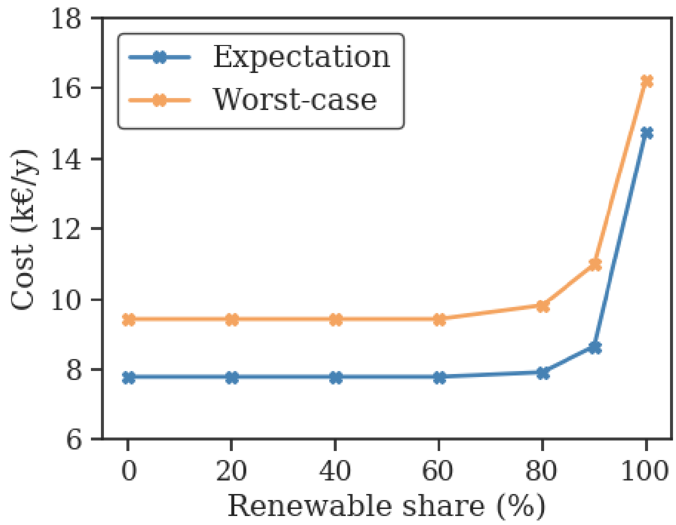


Figure 4.12: Annual system cost for different share of solar production in the risk neutral and worst cases.

Finally, the higher the risk aversion, the higher the equipment sizes and the higher the total system cost. As shown in figure 4.11 and 4.12, the gap is greater when the renewable share is lower than 90%: +20% against +10% for the total cost. This gap is probably due to the fact that the multi-energy system is facing uncertainties from the electricity tariff, the energy demands and the production in the first case, while only the second and third factors remain when autonomy is reached. Furthermore, the size of the thermal storage seems to be the most impacted by the risk measure. So far, the reason for this variation has not been properly understood by the author.

4.6 A closer look at the interaction between design and operation

The objective of this section is to get better insights into the design and operation interaction. Should the operation strategies used for the design (i.e., Φ^d) and in real life (i.e., Φ^a) be the same? This part attempts to answer this question on a simplified case study. In addition, this section quantifies the optimality level of the operation strategy and studies its impact on the design values.

To this end, the DES is simplified to solely focus on this latter methodological aspect rather than the system complexity. The simplified case study is depicted in figure 4.13 where the hydrogen units and the thermal storage are removed. Following the same philosophy, the DES is successively designed and evaluated using several operation strategies with different levels of optimality.

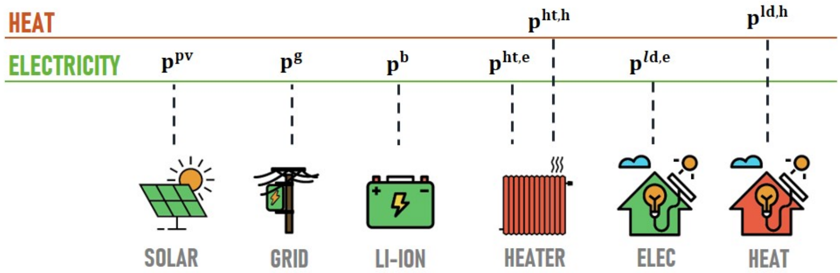


Figure 4.13: Schematic view of the simplified DES.

Therefore, six policies (i.e., 1 anticipative and 5 RBs) with decreasing performances are introduced where the anticipative policy gives the best result. Then, the first rule-based strategy (RB1) is defined as follows: first, the heater is used to supply the thermal demand. Then, the solar production is directly consumed to supply both the electrical demand and the heater consumption. Finally, the battery is charged whenever there is a surplus of energy or discharged otherwise. Then, four other RB policies are built upon the previous strategy by adding an increasing white noise perturbation to the final battery power flow decision. Therefore, the operation strategies range from the most efficient but unrealistic anticipative policy to the last RB strategy with poor performances due to the added perturbation. For the sake of simplicity, only the battery power flow decisions are degraded but more sophisticated perturbation could be added to the controller. The only objective is to end up with a range of strategies that perform differently.

The objective is now to run the metaheuristic design method with the anticipative and noisy RBs successively to compare the resulting equipment sizes. Two special cases are further studied in the following: 1) a first case without any share constraint (i.e. $\tau^{sh} = 0$) where the benefit from installing solar panels and storage systems is purely economic; 2) a second case with a strong renewable share (arbitrarily set to 80%).

4.6.1 Sensitivity of the operation strategy performance over the design

Figure 4.14 shows the sizing as a function of the operation policies. Without any share constraint, the size of the battery decreases to zero along with the performance of the operation policy as it makes no sense to install a battery if it is misused. The anticipative and RB1 policies give approximately the same design results. On the other hand, when 80% of solar production (on average) is required, the optimizer has to oversize the assets to compensate for poor energy management performances: the difference goes up to +70% for solar panels and +40% for the battery between the best and the worst strategy. However, in this second case, design values are less sensitive to the operation policy as the results are nearly the same for the first three strategies.

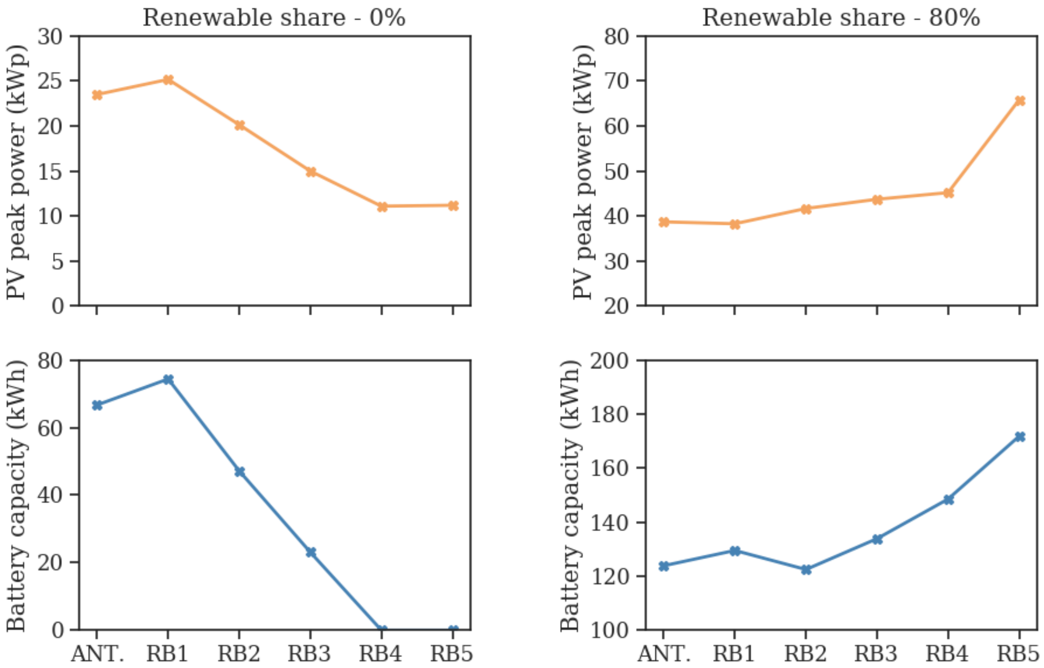


Figure 4.14: Design results for both assets as a function of the operation policy (left) without the share constraint; (right) with a renewable share equal to 80%.

Therefore, the results confirm previous section observations: the operation policy embedded in the design phase Φ^d and the one used in real-life Φ^a do not have to be the same, but their levels of optimality must be similar. In practice, it only means that the mathematical programming method can be used instead of the metaheuristic procedure, only if the system is finally operated in real life with the RB1 or RB2 policies (depending on the renewable share constraint).

4.6.2 Sensitivity of the operation strategy over the out-of-sample cost

Once the system has been designed, the objective of this part is to assess each sizing solution with the different operation strategies. The assessment results over the 1000 scenarios Ω^a are given in figure 4.15. For each table, a row is associated with one design obtained with one of the six corresponding operation strategies. Then, each design is evaluated with all the operation strategies (one for each column) and the resulting total annual cost expected value is printed in the table. When the renewable share expected value is not met, the cell is left blank (figure 4.15b).

		Expected cost (k€/y) - τ^{sh} = 0%					
D	ANT.	8.90	8.90	8.99	9.18	9.38	9.89
E	RB1	8.91	8.91	9.00	9.21	9.42	9.97
S	RB2	8.92	8.91	8.99	9.15	9.31	9.73
I	RB3	8.99	8.99	9.04	9.14	9.25	9.53
G	RB4	9.22	9.22	9.22	9.22	9.22	9.22
N	RB5	9.22	9.22	9.22	9.22	9.22	9.22
		ANT.	RB1	RB2	RB3	RB4	RB5
		ASSESSMENT					

(a) $\tau^{sh} = 0\%$

		Expected cost (k€/y) - τ^{sh} = 80%					
D	ANT.	9.50	9.50	9.64	-	-	-
E	RB1	9.58	9.58	9.72	-	-	-
S	RB2	9.60	9.60	9.74	-	-	-
I	RB3	9.88	9.88	10.02	10.30	-	-
G	RB4	10.24	10.24	10.39	10.68	10.98	-
N	RB5	12.12	12.12	12.27	12.56	12.85	13.59
		ANT.	RB1	RB2	RB3	RB4	RB5
		ASSESSMENT					

(b) $\tau^{sh} = 80\%$

Figure 4.15: Expected total annual cost (k€/y) for each design value evaluated with each operation policy (a) without the share constraint; (b) with a renewable share equal to 80%. When the renewable share expected value is not met, the cell is left blank.

Without any renewable share constraint, the worst total annual costs (about +10% compared to the lowest cost) are obtained when the DES is designed with high-performance policies (anticipative or RB1) and then operated with the low-performance RB5 policy (upper right corner). Note that the cost remains constant when the DES is designed with both RB4 and RB5 (whatever the assessment strategy) as the size of the battery is null (see figure 4.14). Concerning the constrained case, the renewable share might not be met if the operation strategy used to design the DES performs better than the one used for the assessment (blank cells in figure 4.15b). In every case, the total annual cost is lower when the system is finally operated with a more effective policy than the one used in the design phase (lower matrix triangle). Therefore, an important remark from these observations is that **it is safer to use a poor energy management strategy to design the system and then use an operation strategy that performs well in real-time, than the opposite**. Otherwise, techno-economic requirements might not be met, with a higher total annual cost than expected.

But what does "good" or "high-performance" mean for an operation policy? Similarly to [116], a way to quantify the degree of optimality for each strategy is to remind that the anticipative policy gives the best operating cost ($\tilde{J}_s^{o,ant}$), while the worst case is obtained when the battery is unused ($\tilde{J}_s^{o,ref}$) i.e. power flow decisions sent to the battery are zeros. Therefore, a score $q_s \in [0,1]$ is computed for each operation strategy and for each scenario $s \in \Omega^a$ based on the maximum achievable economic gain for a given design solution ($\tilde{J}_s^{o,ref} - \tilde{J}_s^{o,ant}$) (4.15).

$$q_s = \frac{\tilde{J}_s^{o,ref} - \tilde{J}_s^{o}}{\tilde{J}_s^{o,ref} - \tilde{J}_s^{o,ant}} \tag{4.15}$$

where $\tilde{J}_s^{o}$ is the operating cost obtained from the evaluated operation policy.

		Score [0-1] - τ^{sh} = 0%					
D	ANT.	1.00	1.00	0.97	0.90	0.84	0.67
E	RB1	1.00	1.00	0.97	0.91	0.85	0.68
S	RB2	1.00	1.00	0.97	0.89	0.82	0.62
I	RB3	1.00	1.00	0.95	0.86	0.76	0.49
G	RB4	-	-	-	-	-	-
N	RB5	-	-	-	-	-	-
		ANT.	RB1	RB2	RB3	RB4	RB5
		ASSESSMENT					

(a) $\tau^{sh} = 0\%$

		Score [0-1] - τ^{sh} = 80%					
D	ANT.	1.00	1.00	0.97	-	-	-
E	RB1	1.00	1.00	0.97	-	-	-
S	RB2	1.00	1.00	0.97	-	-	-
I	RB3	1.00	1.00	0.97	0.92	-	-
G	RB4	1.00	1.00	0.97	0.92	0.87	-
N	RB5	1.00	1.00	0.98	0.93	0.88	0.76
		ANT.	RB1	RB2	RB3	RB4	RB5
		ASSESSMENT					

(b) $\tau^{sh} = 80\%$

Figure 4.16: Performance score of the different operation strategies for each design value (a) without the share constraint; (b) with a renewable share equal to 80%.

Figure 4.16 shows the expected value of the performance score over the 1000 scenarios Ω^a for both cases (with and without the renewable share constraint). Note that the variance over the 1000 scenarios is lower than 10^{-3} for each value. A first observation is that the anticipative and RB1 policies give the same performance score despite the heuristic nature of the rule-based strategy. Again, note that RB4 and RB5 solutions lead to zero battery capacity so that the performance score does not mean anything in these two latter cases (the cells are left blank in the table). Moreover, the performance metric seems to be approximately constant regardless of the design values for each policy. This remark is not completely true for the RB3 design where the size of the battery is close to zero. Comparing both figure 4.14 and figure 4.16 leads to the following conclusions:

- Design values are more sensitive to the operation strategy performance without the renewable share constraint than in the other case.
- Solving a single large LP optimization problem to design the DES is only relevant whether the performance score of the realistic and anticipative strategies are the same in the first case, or greater than 0.92 in the second case.
- The performance indicator introduced in this section helps quantifying the optimality notion attached to the operation strategy.

4.7 Discussion and conclusions

The optimal design and operation under uncertainties of a multi-energy system with seasonal storage were presented in this chapter. First, the general mathematical formulation was described with a particular focus on the information structure, which is of first importance when dealing with stochastic issues. Next, uncertainties were modeled as discrete random variables over a probability space, and a scenario generation method based on Markov chains was introduced to build the discrete sample space. Then, because the exact solution of such optimization problems is most of the time unreachable, resolution methods for both the design (i.e., mathematical programming and metaheuristic approaches) and operation (i.e., RB, OLFC and anticipative policies) were introduced. These methods have been jointly studied through a modeling framework where the design and assessment phases are divided into two distinct parts. In this way, whatever the assumptions made during the design optimization step, the solutions are assessed on a common out-of-sample simulator with realistic operation strategies which only have access to past and current information.

Results show that the mathematical programming method (widely implemented in the literature) gives the least cost option as long as the DES is finally operated in real life with near-optimal operation policies such as OLFC. Otherwise, the techno-economic requirements might not be satisfied with underestimated design values. In the case where poor performance and fast operation policies (such as RB) are used in real life, the meta-heuristic approach might be a suitable option to provide design values consistent with a given risk measure. Comparing these two approaches allows assessing the potential cost reduction from improving the performance of the operation policy: in this work, the total system cost is reduced by 16% when using OLFC instead of RB. Note that design values are completely different depending on the operation policy embedded into the design procedure. In the last part of this chapter, the value of the stochastic solution was strengthened by its comparison with deterministic designs. As it has been shown, despite shorter computational times, sizing values from deterministic models should be carefully treated as they highly depend on the scenario selected for optimization which could lead to a severe underestimation of the techno-economic system performances.

This work reinforces that the design and the realistic operation of the DES cannot be approached separately. Without this latter part, which is to the best of the authors' knowledge, usually omitted in most of the planning studies, the value of the design solutions might not be relevant for real-world applications. Indeed, most of the papers apply the mathematical programming approach and then run parametric analyses, but they do not discuss or verify that the requirements are met with realistic operation policies. The output techno-economic indicators of the design phase are those of the simplified problem which assumes perfect foresight of the operation. This work provides some insights into the validity of this assumption regarding the realistic operation of the DES. This information might be critical for engineers seeking to install and operate such multi-energy systems in real life.

For the sake of clarity, the same model granularity (e.g., physical models, time step) was used in both the optimization and simulation phases, because the objective of this work was to highlight the relation between the design and realistic operation issues. However, the accuracy of the out-of-sample simulator must be enhanced if quantitative results are needed, with special attention to the uncertainty characterization. These two latter points are definitely future directions for this work to increase the value of the approach for real-life applications. Furthermore, the model implemented in this work is based on a single representative year over multiple scenarios to account for the operation uncertainty. However, in real-world applications, systems are aging and several investment decisions can be made over the horizon of the study. In this case, uncertainty is also added to the design parameters such as the investment costs and long-term evolution of the energy demands. This problem is usually referred to as the *multi-stage planning problem*, and future work needs to be conducted in that direction. This way, in addition to the size of the assets, the optimizer also provides the investment pathway which can be valuable information for decision-makers.

Part III

Toward dynamic design approaches

Chapter 5

Dynamic aware aging design of a simple distributed energy system: a comparative approach with design strategies based on a two-stage model

This is a joint work with Rémy Rigo-Mariani, based on an original mathematical formulation from Rigaut et al [19]

Highlights

- Introduction of a multi-year planning model where the impact of the operation over equipment lifetimes is taken into account.
- This model is compared with two heuristic design strategies based on the two-stage model previously introduced: 1) in the first method, the optimization problem is solved the first year, then systems are replaced with the same sizes at the end of their life; 2) in the second method, the optimization is rerun for each replacement.
- Results show that, unlike heuristic strategies, the dynamic aware aging method can control both the investment pathway and the battery aging, thereby reducing the overall system cost.

Contents

5.1 Introduction

In all previous chapters, design decisions were made based on a single representative year (with multiple scenarios in the stochastic case). The following work deals with the multi-year dynamic design of DES where multiple sizing decisions can be made over the lifetime of the project.

5.1.1 Literature review

The multi-year planning problem has been widely studied in the literature, especially for the long-term evaluation of energy systems at national scales and beyond [32, 42] (usually referred as *expansion planning problems*). The objective of such approaches is to obtain the optimal sizes, location and pathway (i.e., timing of construction and decommission) over the horizon of a project. Until recently, most studies have considered oversimplifications of the system operation with limited snapshots or monthly averages to cope with long computation times caused by the multi-year formulation.

In recent years, the integration of VRE in planning studies has led to the increasing consideration of detailed operation (with accurate temporal and spatial resolution) to properly account for the cost and value of the different technologies at the system scale. For example, Mavromatidis et al [117] developed the MANGO model that incorporates dynamic design for DMES while accounting for detailed operation based on representative days at an hourly time step. The model also includes equipment degradation over years, and multi-location optimization capabilities. Pecenak et al [118] compared two dynamic design strategies for microgrids where the operation is also represented by representative days for each year. Limpens et al [119] used a multi-year model to study different energy system pathways for Belgium. In this case, investment can be made every 5 years and each 5-year block is represented by a single equivalent year.

The studies above were derived in a deterministic framework, and the operation is mostly based on representative periods to tackle long computation times. The reason for this is that the multi-year dimension significantly increases the model complexity, leading to inevitable simplifications of the other facets of the original problem. However, the nature of the multi-year problem is highly stochastic because it is unlikely that the future parameters (e.g., investment and energy costs, energy demands) remain unchanged over time. As a result, a growing number of studies have sought to include both short- and long-term uncertainties in the multi-year planning process [44, 45, 46, 120, 121, 122]. In this case, the resolution of such complex optimization problems usually requires the development of new mathematical methods (mostly based on decomposition techniques).

All the previous references assume that the assets have fixed lifetimes as it facilitates mathematical implementation. Indeed, because the asset lifetime is known *a priori* and does not depend on the operation, the decommissioning time is also known in advance. However, the interaction between the design and the operation is not fully captured as the way systems are operated has no consequences on their lifetime. Even if some studies include technology performance degradation (for example [117] and [118]), these degradation do not impact the lifetime of the equipment. This latter hypothesis might be strongly discussed for DES including battery storages because, in addition to calendar aging, the number of charge/discharge cycles has great consequences on the equipment lifetime [123, 124].

Unlike previous studies, the authors in [62] and [125] attempted to consider the impact of the operation over the design in the planning process. However, in their case, they only optimized the size of the assets the first year (i.e., design decisions are single values). Then replacements are taken into account by simulating the operation over the entire horizon and systems are replaced with the same sizes. In this case, the multi-year dimension is not included in the design procedure i.e. replacements are not decisions of the optimization problem. Hence, the optimizer does not provide any information about the optimal investment pathway. These approaches have nothing to do with the previous references where the optimizer can decide whether or not to install (or decommission) technology at each design time step.

5.1.2 Main contributions

This work addresses the aforementioned shortcomings by bridging the gap between the different time scales which is deemed necessary to represent the DES lifetime in the design phase, from hourly variations to long-term forecasts of energy prices and cost of technologies. The main challenge tackled by this work is to investigate to what extent the design performance is increased when aging is taken into account in the DES design phase compared to standard methods based on a single equivalent year. To address this question, a multi-time scale model is formulated in a deterministic framework for the toy-example introduced in chapter 1 (see figure 5.1). The DES remains intentionally simple as the objective is to focus on the methodology and the novelties given by this approach rather than the system complexity. Applying the proposed methodology to more complex DES is definitely the intended direction of future works. Then, two heuristic design strategies based on a single equivalent year are compared to the aware aging design method on a common simulation referential which includes the battery aging model and the replacement dynamic properly modeled. The comparison is run at an hourly time step over 20 years (see figure 5.2). In that way, the performance of each approach could be rigorously assessed, despite the simplification made for the purpose of optimization. The objective is to provide a common framework, not commonly encountered in the literature, to equally discriminate different planning strategies.

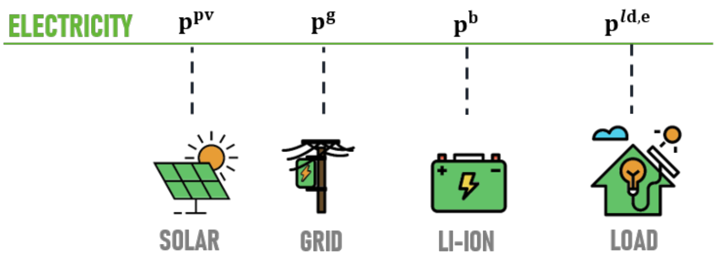


Figure 5.1: Schematic view of the DES with solar panels and a battery.

This work is organized as follows: section 5.2 shows the problem formulation of the multi-time scale model and the optimization problem statement. Then, section 5.3 describes the resolution methods which are going to be compared. Finally, results are shown in section 5.4, conclusions and perspectives are drawn in section 5.5.

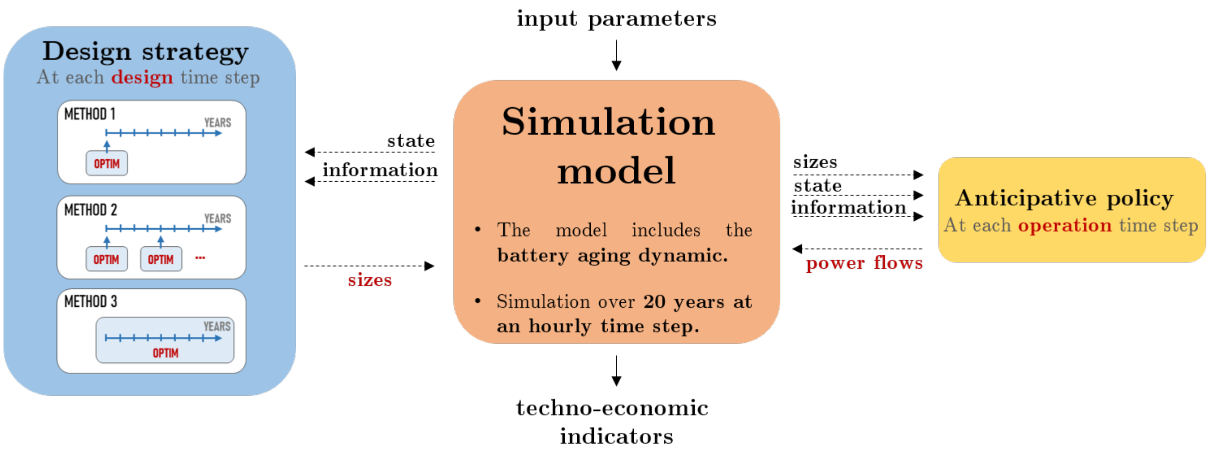


Figure 5.2: Three design methods are compared (depicted in the blue rectangle): they give the design decisions to be made each year based on the current state of the system and available information. The different approaches are evaluated on a common simulation referential which includes the battery aging and the replacement dynamic properly modeled. The comparison is run at an hourly time step over 20 years.

5.2 Mathematical formulation

This section describes the fundamental mathematical equations that allow modeling the DES operation and the impact of design choices. The simulator referential used to evaluate the different design strategies is based on this formulation.

The formulation is then based on Rigaut et al [19] applied to the design and operation of the DES. Both the design and operation dynamics are included in a common optimization problem as described in the following. The contribution of this chapter is towards the comparison of the aware-aging model with other design strategies depicted in section 5.3. To this end, the current study is first limited to the deterministic case. This latter hypothesis is further discussed in section 5.5.

5.2.1 Notations

The design dynamic is a slow process compared to the operation where power flow decisions need to be made every hour. Hence, we define two time scales with the set of years $y \in \mathbb{Y} = \{1, \ldots, Y\}$ and hours $h \in \mathbb{H} = \{1, \ldots, H\}$ with Δy and Δh the two time steps respectively. A time continuum is ensured as depicted in figure 5.3. design decisions are made along the set $\mathbb{Y}$ (orange nodes) while the operations along both sets $\mathbb{Y}$ and $\mathbb{H}$ (blue nodes).

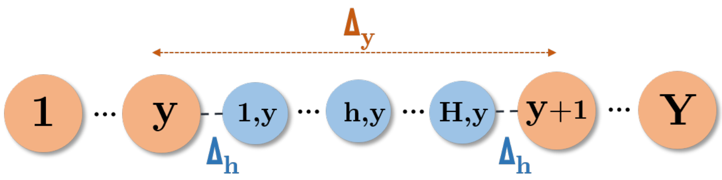


Figure 5.3: A time continuum is ensured in the simulation between the design (y) and operating (h) time scale

5.2.2 Decision variables

The decision variables for both the design and operation optimization problems are given as follows:

- The decision variables for the design are the new sizes of the assets that could be made every year. They correspond to the battery maximum capacity and the PV peak power, gathered in the vector $u_y^d = (E_y^{b,d} \quad p_y^{pv,d}) \in \mathbb{U}_y^d = \mathbb{R}^2$. Contrary to previous chapters, the design decisions are no longer single values.

- The decision variables for the operation are the power flows controlled in the system at every hour. They correspond to the charging and discharging power for the battery, gathered in the vector $u_{h,y}^o = (p_{h,y}^{b,+} \quad p_{h,y}^{b,-}) \in \mathbb{U}_{h,y}^o = \mathbb{R}^2$.

5.2.3 Constraints

The constraints are distinctly introduced for both the design and the operation time scales. Then, the last part shows how to bridge the gap between the two time scales.

Design time scale

Unlike previous chapters, state variables are also introduced for the design to model the replacement dynamic due to aging. They represent the sizes of existing technologies that need to be updated based on the value of the design decisions u_y^d each year. They are denoted by $E_y^{b,state}$ (in kWh) and $p_y^{pv,state}$ (in kWp) for the battery and PV, respectively.

The sizes of the existing equipment could be increased or downscaled depending on the case study. Note that systems are assumed to be completely replaced when a design decision is made. In this case, the dimensions of the existing system $E_y^{b,state}$ and $p_y^{pv,state}$ need to be updated with the new installed capacities. Otherwise, they remain the same as the year before. This design dynamic is given by (5.1) and (5.2). Because solar panels aging is not considered in the current study, a linear formulation would have been possible for the design dynamic of the PV. As it has little consequences on computational times, this generic formulation was adopted in this work.

$$p_{y+1}^{pv,state} = \begin{cases} p_y^{pv,d} & \text{, if } p_y^{pv,d} > 0 \\ p_y^{pv,state} & \text{, otherwise} \end{cases} \tag{5.1}$$

$$E_{y+1}^{b,state} = \begin{cases} E_y^{b,d} & \text{, if } E_y^{b,d} > 0 \\ E_y^{b,state} & \text{, otherwise} \end{cases} \tag{5.2}$$

The design decisions are both continuous, positive and bounded variables (5.3), (5.4).

$$0 \leq p_y^{pv,d} \leq \overline{p}^{pv,d} \tag{5.3}$$

$$0 \leq E_y^{b,d} \leq \overline{E}^{b,d} \tag{5.4}$$

Operation time scale

The operation constraints are almost the same as the toy-example introduced in chapter 2. The only difference comes from the aging equation which is added to the "energy system model" constraints.

To this end, the battery state-of-health (SoH) $A_{h,y}^b$ in hour h and year y - expressed in kWh - is introduced to take the lifetime of the battery into account over the horizon. The aging dynamic is computed using a simple model (5.5) based on the maximum exchangeable energy during its lifespan [123]. This model only considers aging due to cycling. Calendar aging or degradation of the battery parameters (e.g., loss of capacity and increased internal resistance over time) are not taken into account. As in [123], the maximum exchangeable energy depends on the battery maximum number of cycles n_c for a fixed depth-of-discharge dod (5.6). These parameters are usually provided by constructors for each battery technology.

$$A_{h+1,y}^b = A_{h,y}^b - (p_{h,y}^{b,+} + p_{h,y}^{b,-}) \cdot \Delta h \tag{5.5}$$

$$0 \leq A_{h,y}^b \leq 2 \cdot n_c \cdot dod \cdot E_y^{b,state} \tag{5.6}$$

Bridging the gap between time-scales

In order to ensure the inter-year continuity for the battery SoC and the SoH, the decision process is defined as follows: design decisions are made at the end of the last hour of each year (H, y) when the SoH and SoC are completely known over the current year. The sizes of the existing assets are then updated with the new sizes at the beginning of the first hour of the next year $(1, y+1)$. Furthermore, any newly installed battery is assumed to be fully charged and with maximum exchangeable energy (i.e. maximum SoH). Thus, continuity equations for the SoH and SoC between years are given by (5.7) and (5.8).

$$E^b_{1,y+1} = \begin{cases} \overline{e} \cdot E^{b,d}_y & \text{, if } E^{b,d}_y > 0 \\ E^b_{H+1,y} & \text{, otherwise} \end{cases} \tag{5.7}$$

$$A^b_{1,y+1} = \begin{cases} 2 \cdot n_c \cdot dod \cdot E^{b,d}_y & \text{, if } E^{b,d}_y > 0 \\ A^b_{H+1,y} & \text{, otherwise} \end{cases} \tag{5.8}$$

The problem state variables for each year y are gathered in vector $x_y \in \mathbb{X}_y$ (5.9). It includes the size of the assets along with all the SoC and SoH values over the year.

$$x_y = (E^b_{1:H+1,y} \quad A^b_{1:H+1,y} \quad E^{b,state}_y \quad p^{pv,state}_y) \tag{5.9}$$

5.2.4 Optimization problem statement

The objective is to determine both the sizing and operation decisions in order to minimize the sum of both the discounted design and operating expenditures over the 20-years horizon.

Investment cost

The annual capital cost depends on the design decisions u^d_y, and the capital cost of each technology (5.10).

$$J^d_y(u^d_y) = c^b_y \cdot E^{b,d}_y + c^{pv}_y \cdot p^{pv,d}_y \tag{5.10}$$

where c^{pv}_y and c^b_y are the investment costs of solar panels (€/kWc) and the battery (€/kWh) for each year, respectively.

Operating cost

The operating cost is computed based on the energy exchanged with the external network (5.11).

$$J^o_{h,y}(u^d_y, u^o_{h,y}) = (c^{g,+}_{h,y} \cdot p^{g,+}_{h,y} - c^{g,-}_{h,y} \cdot p^{g,-}_{h,y}) \cdot \Delta h \tag{5.11}$$

where $c^{g,+}_{h,y}$ is the tariff of electricity (€/kWh) and $c^{g,-}_{h,y}$ the feed-in tariff (€/kWh), set to zero in the rest of the study. Note that those rates could change over the horizon as they are also indexed by y which motivates the proposed multi-time scale approach.

Salvage value

A salvage value is also introduced to account for the remaining life of the battery at the end of the horizon Y. Its value is computed according to [126] which assumes a linear depreciation of components over time.

$$J^s(x_Y) = K \cdot A^b_{H+1,Y} \tag{5.12}$$

The salvage coefficient K is the ratio of the discounted investment cost over the maximum exchangeable energy at the end of the horizon (5.13).

$$K = \frac{\gamma_Y \cdot c^b_Y \cdot E^{b,state}_Y}{2 \cdot n_c \cdot dod \cdot E^{b,state}_Y} = \frac{\gamma_Y \cdot c^b_Y}{2 \cdot n_c \cdot dod} \tag{5.13}$$

The value of the discount factor γ_y for each year is given by (5.14).

$$\gamma_y = \frac{1}{(1+\tau)^y} \tag{5.14}$$

where τ is the discount rate, set to 4.5% in this study.

Problem statement

The formulated problem aims at finding the optimal decision variables for both the operation and the design in order to minimize the total cost of the system over the horizon (5.15).

$$\min_{u^d, u^o} \sum_{y=1}^{Y} \gamma_y \left[J^d_y(u^d_y) + \sum_{h=1}^{H} J^o_{h,y}(u^d_y, u^o_{h,y}) \right] - J^s(x_Y) \tag{5.15a}$$

s.t.

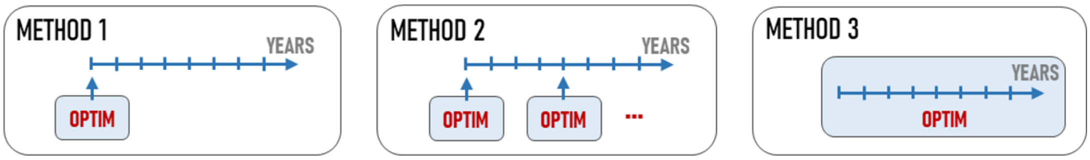
$$x_{y+1} = f(x_y, u^d_y, u^o_{1:H,y}) \tag{5.15b}$$

$$u^d_y \in U^d_y, \quad u^o_{1:H,y} \in U^o_{1:H,y}(u^d_y, x_y) \tag{5.15c}$$

where f is described by the operation and design dynamic equations previously introduced.

5.3 Resolution methods

Problem (5.15) can be solved using traditional solvers but its resolution is computationally intensive. The objective of this section is to introduce two heuristic strategies derived from the two-stage model (described in chapter 4) to solve the multi-year design problem in a shorter time. The results from these two heuristics will be compared to the exact solution of the problem in section 5.4. The different design approaches are depicted in figure 5.4.



(a) (b) (c)

Figure 5.4: Three design methods are compared in the next section. (a) and (b) are two heuristic design strategies derived from the two-stage model introduced in chapter 4. In the first case, the problem is only solved once at the beginning of the horizon, then the battery is always replaced with the same capacity when it reaches its lifetime during simulation. In the second method, the design optimization is rerun every time the battery reaches its end of life over the horizon. The last method (c) finds the exact solutions of the multi-year problem.

5.3.1 Method 1: two-stage model based on a single representative year

The first heuristic method is derived from the two-stage model introduced in chapter 4 applied to the toy-example. Each year y of energy demand and production (i.e., $\{1, \ldots, Y\}$ in total) corresponds to a scenario s in the two-stage formulation of section 4. The expectation risk measure is used for the total annualized cost, while the renewable share constraint must be met at all times.

Therefore, the design strategy aims at solving the problem in the first year. Then, the battery is always replaced with the same capacity when it reaches its lifetime during simulation. The resulting design strategy ϕ^1 is then given by (5.16) which depends on the battery SoH at each design time step.

$$\phi^1_y(x_y) = \begin{cases} u^{d*} = \underset{u^d, u^o}{\arg\min} \text{ (problem 4.9)} & , \quad \text{if } y = 1 \\ u^{d*} & , \quad \text{if } A^b_{H+1,y} \leq \epsilon \\ 0 & , \quad \text{otherwise} \end{cases} \tag{5.16}$$

where ϵ is a parameter fixed by the user. In this work, the battery is replaced when the SoH is lower than 10% of its initial value.

5.3.2 Method 2: two-stage model based on a single representative year with online re-optimization

The second method is similar to method 1 except that the two-stage optimization model is rerun every time the battery reaches its end of life over the horizon. Instead of replacing the battery with the same capacity computed at the first year, new design decisions can be made with updated information (e.g. equipment cost, energy prices). Because the PV lifetime is assumed to be longer than the horizon, its capacity remains fixed to its first-year value and no replacement may be needed. The formulation is unchanged from method 1 but the investment costs and the number of remaining years Y before the end of the horizon are updated. The heuristic design strategy ϕ^2 is given by (5.17).

$$\phi_y^2(x_y) = \begin{cases} u^{d*} = \underset{u^d,u^o}{\arg\min}\ (\text{problem 4.9}) & , \quad \text{if } A^b_{H+1,y} \leq \epsilon \\ 0 & , \quad \text{otherwise} \end{cases} \tag{5.17}$$

5.3.3 Method 3: aware aging design based on the multi-time scale model

The third method finds the exact solution of problem (5.15) by formulating a single large MILP problem. Big-M values with binary variables are introduced in order to linearize "if-else" functions (5.2) - (5.8) as it is commonly done in MILP formulation. As an example, the replacement dynamic equation (5.2) for the battery becomes (5.18) and (5.19) and allows ensuring the inter-year continuity as previously mentioned.

$$-M \cdot (1 - \delta_y^b) \leq E_{y+1}^{b,state} - E_y^{b,d} \leq M \cdot (1 - \delta_y^b) \tag{5.18}$$

$$-M \cdot \delta_y^b \leq E_{y+1}^{b,state} - E_y^{b,state} \leq M \cdot \delta_y^b \tag{5.19}$$

M is the big-M value equal to the sizing bounds and δ_y^b a binary variable which is equal to 1 when the battery has to be replaced (i.e., $E_y^{b,d} > 0$). In that way, if $E_y^{b,d} > 0$ then $E_{y+1}^{b,state}$ is equal to $E_y^{b,d}$ thanks to (5.18), otherwise the capacity is unchanged from the previous year (5.19).

Note that since the problem has been modeled in a deterministic framework, the three resolution methods provide both the design and operation decisions at the same time (similarly to chapter 2). In other words, the operation policies that will be used to evaluate the different design approaches in the next section are, by construction, anticipative.

5.4 Numerical results

This section aims at demonstrating novelties from the multi-time scale approach compared to the other resolution methods. Several examples will be introduced in order to highlight some specific points.

5.4.1 Case study

Input parameters [1] are listed below:

- The electrical demand and production profiles come from Ausgrid. For each year of the 20-years horizon, 10 consumers profiles were randomly chosen and aggregated as input of the study. The demand and production variability between years is taken into account as each yearly profile is different from the others.

- The tariff of electricity follows a peak/off-peak EDF ”Tarif Bleu” [127]. For the sake of simplicity, no price evolution is taken into account over the 20 years.

- The cost of storage system (Li-ion battery + converter) is given by [128] and decreases from 600 €/kWh in 2021 to 300 €/kWh in 2040.

- The cost of solar panels including AC/DC converters is given by [129] and decreases from 1040 €/kWc in 2021 to 735 €/kWc in 2040.

- Technical parameters for the battery are reported in table 5.1.

In what follows, the net present value (NPV) is also computed along with the total cost over the horizon, to assess the performance of each method. It is a standard metric to compare different investment projects [126, 130] as it gives extra information about the profitability along the system lifetime. This latter quantity is computed from the difference between positive cash flows and investment costs. A positive NPV results in profit: the higher the NPV at the end of the horizon, the higher the profitability. In this work, cash flows correspond to savings compared to the baseline cost where all the electricity is purchased from the grid.

Parameters	η^- [0-1]	η^+ [0-1]	$\underline{e}$ [0-1]	$\overline{e}$ [0-1]	$\underline{p}$ [h^{-1}]	$\overline{p}$ [h^{-1}]	n_c [-]
Battery	0.8	0.8	0.2	0.8	1.5	1.5	2500

Table 5.1: Battery parameters

The problem is modeled using Julia and JuMP package [72]. The IBM CPLEX 12.10 solver is then used to solve the problem. All the computations are run on a Intel Xeon(R) CPU E5-2697 v2 @ 2.70 GHz x 48 server.

[1] It would have been more relevant to apply the methodology on a case study similar to chapter 4. However, this study was conducted before chapter 4, and it was decided to apply the stochastic approaches to a different case study. Hence, other parameters (e.g., battery efficiency, investment costs) are also different from those of the previous chapter.

5.4.2 Example 1: comparison with a renewable share constraint set to 60%

With the current low electricity tariffs and costs of technologies, none of the three methods would invest in a Li-ion battery due to high capital expenditure compared to the expected savings on the operation. Thus, to compare the results from the different approaches, a first example is introduced where the renewable share ratio is arbitrarily set to 60% in order to justify the installation of a storage system.

Figure 5.5 shows the planning strategy over the horizon for the three methods with a 60% renewable share ratio. Note that the assets are assumed to be installed at the end of the first year in order to be operated the first hour of the second year. As shown in the figure, method 3 installs 75 kWp of solar panels and 182 kWh of battery in the first year. Then, the battery is replaced once the 9th year by a new 205 kWh asset (+23 kWh compared to the initial investment) which occurs when the SoH reaches zero (see figure 5.6). Figure 5.6 shows that the multi-time scale formulation handles correctly continuity issues and aging is controlled in order to install the new battery when the previous is out of order.

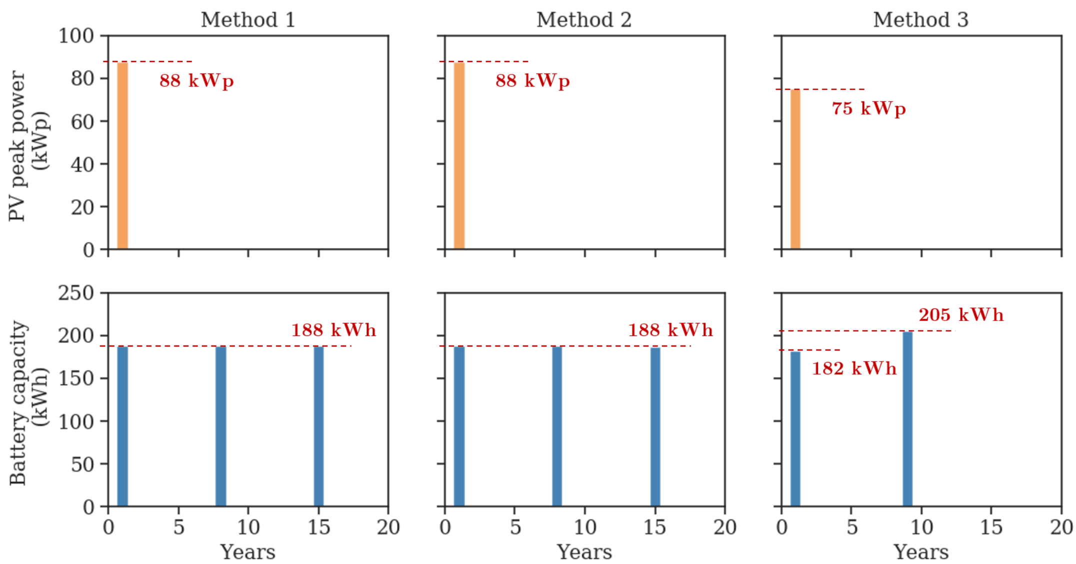


Figure 5.5: Planning strategy over the horizon with a 60% renewable share ratio.

For methods 1 and 2, the sizing results are the same in both cases with 88 kWp of PV and 188 kWh of storage installed the first year. Then the battery is replaced two times, the 8th and the 15th year with the same capacity as in the first year. Remember that method 2 reruns the optimization at the end of the battery lifetime whereas method 1 installs the same initial capacity at each replacement. The optimizer in method 2 does not take advantage of the investment cost reduction to increase the battery capacity when it has to be replaced. Indeed, even with cost decrease and current energy prices, it is not economically profitable to purchase battery storage, so the optimizer installs just enough battery and PV to ensure the renewable share constraint fulfillment until the end of the simulated horizon. In both cases, the SoH is not controlled during operation and the battery needs to be replaced one more time compared to method 3. However, at the end of the horizon, the storage systems are still ”usable” and the remaining SoH (as a percentage of the total exchangeable energy) is 35% for both methods.

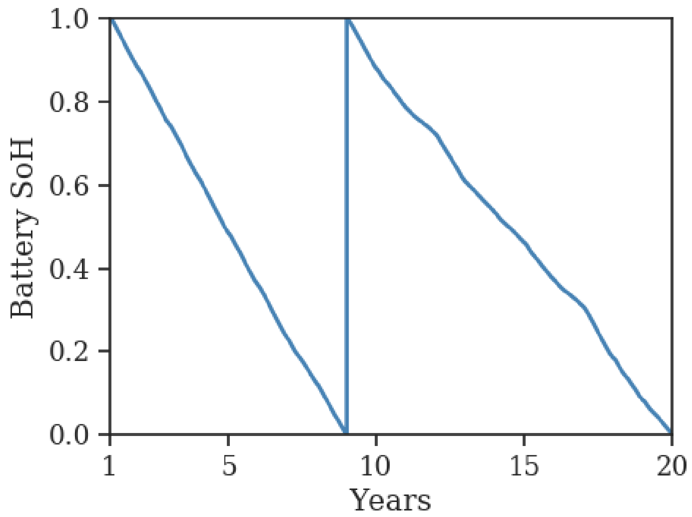


Figure 5.6: The battery SoH over the horizon with method 3 and a 60% renewable share ratio. The SoH has been normalized by the maximum exchangeable energy in the figure.

As shown in figure 5.7, the renewable share constraint is fulfilled each year with every method. Concerning method 3, the optimizer operates the DES and installs just enough capacity to fulfill the constraint each year with less margin than with the other strategies. In contrast, methods 1 and 2 optimize the system to fulfill the constraint during the worst years when the net energy demand is the highest which leads to a greater renewable share ratio for the rest of the years as the operation takes advantage of the system oversizing.

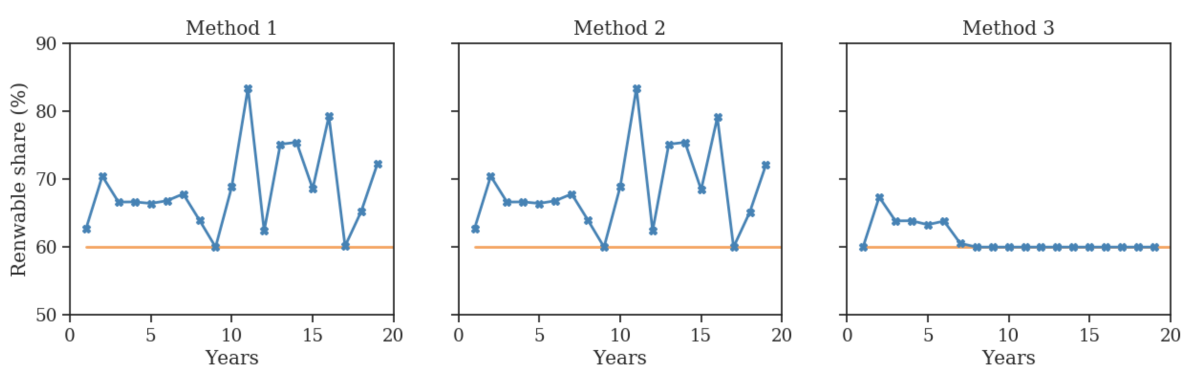


Figure 5.7: Renewable share constraint for each method after simulation. Its value was set to 60%.

Economical results are shown in table 5.2. Note that with the current economic hypothesis, the DES is not profitable over 20 years because of the high installation cost of the battery. The reference cost is then lower than the optimal results obtained with the renewable share constraint and the NPV is here given for information purposes only (negative values for all the methods). As shown in the table, the total discounted cost over the 20 years is 10% lower with method 3 than with the other methods. Note that the results of methods 1 and 2 are approximately the same.

	Ref.	Method 1	Method 2	Method 3
OPEX [k€]	185.8	62.5	62.5	72.4
CAPEX [k€]				
Li-ion	-	205.4	205.2	167.9
PV	-	87.0	87.0	74.8
Salvage [k€]	-	-8.4	-8.3	-
Total [k€]	**185.8**	**346.5**	**346.4**	**315.1**
NPV [k€]	-	-160.7	-160.7	-129.3
CPU time	-	5 min	3×5 min	24 h

Table 5.2: Comparison of the total discounted costs over the 20 years and CPU time (computational time required to solve the model) between the three methods with a renewable share ratio set to 60%. The reference case where all the electricity is bought from the grid and the NPV are given for information purposes.

Finally, the better performances of method 3 are obtained at the cost of greater computational time with 24h long calculation where method 1 is run in only 5 min and method 2 in 15 min. Even if the number of variables and constraints are in the same range between the different methods (see table 5.3), the complexity in the third approach is increased with the introduction of binary variables for the inter-year continuity on the investment decisions.

	Variables	Constraints
Method 1 & 2	876 022	1 927 304
Method 3	1 051 362 (38 bin)	2 444 637

Table 5.3: Comparison of the complexity between the three methods.

As a conclusion of this section, the results could be summed up as follows:

- With the multi-time scale approach, aging and number of replacement are controlled over the whole time horizon.
- The sizing could be increased or down-scaled each year with method 3.
- It pays to control aging as the total discounted cost over the horizon is 10% lower than with the other strategies on the considered test case.

- The last method suffers from longer computational times compared to the other approaches.

5.4.3 Example 2: comparison with a higher electricity tariff

The aim of this section is to compare the three methods in a case where it is profitable to install a Li-ion battery without the renewable share constraint. This could be obtained by assuming a lower investment cost of systems or with increased tariffs of electricity. The latter strategy is chosen and the tariff of electricity is multiplied by 5 in order to have a battery installed the first year.

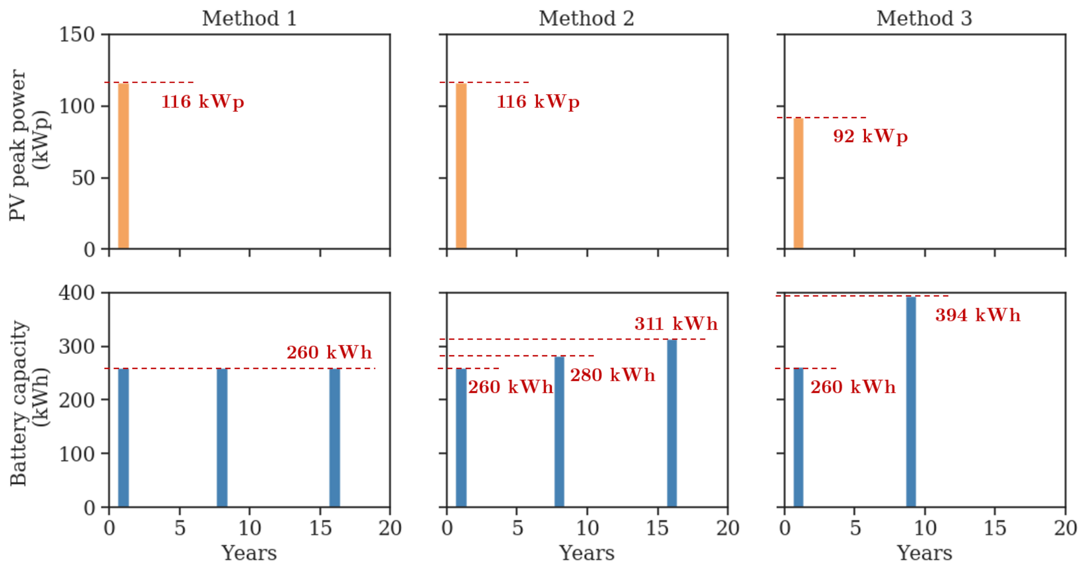


Figure 5.8: Planning strategy over the horizon with a tariff of electricity multiplied by 5.

As shown in figure 5.8, every method approximately installs the same battery capacity the first year (260 kWh). Then, both methods 2 and 3 take advantage of lower storage costs to increase the size of the battery. While method 3 only replaces the battery once with a 394 kWh asset (+134 kWh), method 2 replaces the storage twice with greater capacities: 280 kWh (+20kWh) the first time and 311 kWh (+51kWh) the second time. Similar to the test performed in the previous section, the battery is out of order at the end of the horizon for method 3 whereas it remains 50% and 54% of the total exchangeable energy for methods 1 and 2 respectively. Concerning the PV, both methods 1 and 2 install 116 kWp of solar panels which is 24 kWp greater than method 3. Compared to the previous section, the solar panel size increases when the electricity gets more costly, but no additional investment is made over the horizon. This latter aspect may be explained because when investment decisions are made, technologies are entirely replaced by new systems according to the modeling developed in this work. The cost induced is then proportional to new sizes instead of additional investment only. Thus, technologies with longer lifespans than the horizon are preferably installed in the first year. Similar to previous section, the economic results are given in table 5.4. As previously mentioned, the DES is profitable in every case and the NPV at the end of the 20 years corresponds to profits compared to the baseline case without solar panels and batteries. The differences between the NPV values are lower than in the previous case: the profit is 4% greater with method 3 than method 1 and 3.6% greater than method 2. Note that the resulting NPV of methods 1 and 2 are again approximately the same.

	Ref.	Method 1	Method 2	Method 3
OPEX [k€]	929.0	210.5	197.6	213.0
CAPEX [k€]				
Li-ion	-	280.3	296.8	270.8
PV	-	115.8	115.8	91.8
Salvage [k€]	-	-16.6	-21.9	-
Total [k€]	**929.0**	**590.0**	**588.3**	**575.6**
NPV [k€]	-	339.1	340.8	353.4
CPU time	-	5 min	3×5 min	24 h

Table 5.4: Comparison of the total discounted costs over the 20 years and CPU time between the three methods with the tariff of electricity multiplied by 5.

An interesting point is highlighted in figure 5.9. In that case, the tariff of electricity is multiplied by 3. Unlike the other methods, method 3 installs 48 kWp of solar panels the first year but waits until the 11th year to purchase a 134 kWh battery when its cost is divided by 1.5. The multi-time scale approach not only optimizes the size of technologies but also determines the optimal timing (investment pathway) to install the components along the horizon. This result is only made possible with the use of the proposed method 3 as the design is reconsidered each year in the formulation. This latter aspect gives more flexibility to the approach when strong evolution occurs over the input parameters.

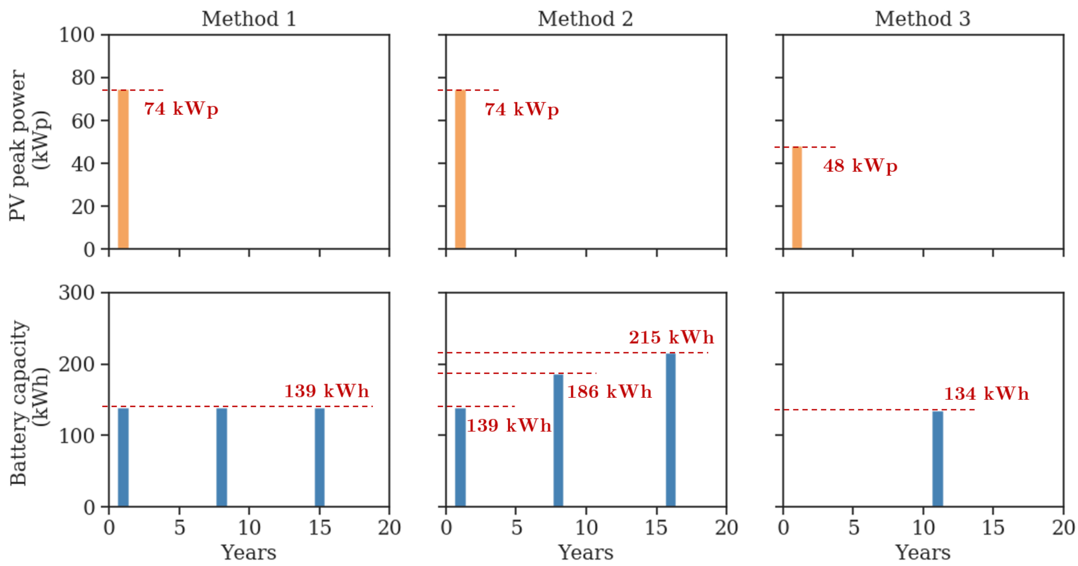


Figure 5.9: Planning strategy over the horizon with a tariff of electricity multiplied by 3. Unlike the other methods, method 3 waits until the 11th year to install the battery at a lower cost than first year.

Again, the NPV of methods 1 and 2 are approximately the same while it is 20% higher for method 3 in this case.

	Ref.	Method 1	Method 2	Method 3
OPEX [k€]	557.4	233.7	214.3	345.7
CAPEX [k€]				
Li-ion	-	153.2	179.8	35.3
PV	-	74.2	74.2	48.1
Salvage [k€]	-	-6.2	-14.5	-
Total [k€]	**557.4**	**454.9**	**453.8**	**429.1**
NPV [k€]	-	102.4	103.5	128.4

Table 5.5: Comparison of the total discounted costs over the 20 years between the three methods with the tariff of electricity multiplied by 3.

5.5 Discussion and conclusions

A generic multi-time scale model for both the design and operation of energy systems was presented in this work. The integration of the aging and replacement dynamic into MILP formulation was depicted. Two examples were run to demonstrate novelties that could be extracted from this approach. Results were then compared to more standard approaches which consider a single equivalent year for the optimization of the design.

It is shown from the previous sections that it pays to control aging as the optimal solutions from the proposed multi-time scale model give the best economic results: the difference goes up to 20% of the total cost in some cases. This cost improvement is made possible because the multi-time scale model includes the interaction between design and operation with investment options every year. Indeed, the aging of the battery and thus, the number of replacements is controlled over the horizon, which is not possible with methods 1 and 2. Moreover, method 3 gives more flexibility to the design of the DES as it can accommodate input parameters evolution over time: the size of systems could be modified along the horizon by taking full advantage of the new information available. Furthermore, more than only optimizing the size of systems, the multi-time scale approach also determines the investment strategy over the horizon. This latter point seems to be interesting for industrial who are willing to know when would be the best moment to invest in order to maximize profitability. Again, this feature could not be addressed by standard methods.

Unlike [45] and [131], the multi-time scale formulation focuses on the interaction between the investment and the operation which is, to the best of the authors' knowledge, rarely addressed in the literature. For the sake of simplicity, equipment were assumed to be entirely replaced with the modeling developed in this work and no additional investment is made possible. This latter point is definitely a weakness of the model and future work needs to be conducted in that direction to increase the value of the approach. Furthermore, the study was conducted in a deterministic framework whereas the real problem is profoundly stochastic: investment and operation decisions have to be made without perfectly knowing the future. However, the aim of the study was to make a first evaluation in a deterministic framework to assess the cost reduction potential of the multi-time scale approach before going into more realistic modeling with uncertainties. In order to include the stochastic nature of the problem and apply the methodology to more complex energy systems, computational times need to be reduced and further resolution methods have to be explored (the better performances of method 3 are obtained at the cost of greater computational time with 24h long calculation where method 1 is run in only 5 min and method 2 in 15 min). For instance, the decomposition methods developed by [19] seems to be a great candidate to address this challenge.

Despite these aforementioned limitations, it seems that the multi-time scale approach could be a promising direction in order to improve the income of a given project when grid rates are low and standard strategies fail to find out a profitable design. The model has proven to be a relevant approach and this work could be seen as a starting point for whoever would be interested in the integration of the interaction between investment and operation in more general energy models and complex analysis.

Part IV

Software

Chapter 6

Genesys.jl: a generic toolbox to assess and compare multiple design and operation approaches

The code is freely available at `https://github.com/hradet/Genesys.jl`.

Highlights

- A generic Julia toolbox to assess and compare different design and operation strategies without changing the overall code.
- An example is given where the two-stage problem is solved for the toy-example, and the results are assessed using a rule-based policy.
- Custom design and operation approaches can be easily added to the framework.

Contents

6.1 Introduction

Coding is an essential part of this work as appropriate tools are necessary to get valuable results. Along this PhD journey, a significant number of hours were spent in front of this black box Atom [132] screen, thinking about the code architecture, its potential reuse, and writing a multitude of code lines. However, despite its critical importance, software implementation is usually less valued than the results. This section is added to this manuscript to pay tribute to the hard coding work, as it constitutes the foundation of all the previous results shown in this thesis.

A wide range of planning and simulation tools have been already developed in the literature to deal with the design and operation issues of energy systems [14, 133]. As previously said in chapter 1, a difference is usually made between *simulation* and *optimization* tools, leading to different scientific communities with little overlap [14, 16]. On the one hand, simulation tools (e.g., EnergyPlan [134]) are generally used to simulate energy system operations with fixed architecture, allowing ”what-if” analyses by running the model several times with different input parameters. On the other hand, optimization tools (e.g., OSeMOSYS [135], PyPSA [136], Calliope [137]) are mainly developed to optimize the design of energy systems, depending on case study requirements. In most cases, they can be used at multiple scales, ranging from buildings to entire continents. Also, a growing number of modeling tools include dynamic investment optimization issues, as previously introduced in chapter 5.

The main drawback of these ”ready-made” tools is that single design and/or operation methods are usually implemented in the software package as their main objective is to come up with optimized solutions. While the majority of software packages handle sensitivity analysis, the latter is only run over the input parameters. However, the sensitivity of the design and operation methodologies over the results cannot be assessed, especially in a coordinated manner. As a result, the software architectures have not been designed in a generic way to handle multiple sizing and operation strategies without changing the overall code.

To address these shortcomings, Genesys.jl has been developed as a generic Julia [138] toolbox to assess and compare different design and operation approaches for energy systems. The objective is to provide a unique and straightforward framework for both the design and the evaluation of the sizing solutions with different operation strategies. The toolbox enables easy integration of custom design and operation methodologies without modifying the entire code structure. This feature is made possible thanks to the ”multiple dispatch” capability of the Julia language. Also, the toolbox addresses simulation over multiple years (i.e., dynamic investment) and multiple scenarios (i.e., stochastic) at an hourly time step. Similarly to the PyPSA philosophy, the toolbox has no graphical interface and the user interacts through Jupyter notebooks or any other Julia IDE (e.g., Juno)

The chapter is organized as follows: first, an overview of the toolbox core ingredients is given in section 6.2. Then, section 6.3 shows an application where the two-stage model is solved for the toy-example. Next, section 6.4 describes how to use custom design and operation strategies within this framework. And finally, the conclusion and perspectives are drawn in section 6.5.

6.2 Toolbox description

As previously said in the introduction, the toolbox enables simulation over multiple scenarios and years at an hourly time step. The time scale is split into both hours and years to account for decisions at the different dynamics: design decisions are made at yearly time steps, while operation decisions are made at hourly time steps. In practice, it only means that operation variables have 3 dimensions (i.e., one for each time set (hours and years), and the last dimension for scenarios), while design variables only have 2 dimensions (i.e., years and scenarios). Continuity issues between two consecutive years are addressed following the formulation introduced in chapter 5.

6.2.1 Structure overview

The toolbox is made of 4 essential components, defined as individual Julia composite type (also called *structure*):

- **Microgrid**[1]: the "Microgrid" structure (in listing 6.1) is the key ingredient of a simulation, and it must be first initialized before going any further. The global parameters (e.g., time steps, hourly and yearly horizons, number of scenarios, renewable share, discount rate) are stored in the microgrid "parameters" field. Furthermore, the assets are gathered in specific vectors depending on their type: demands, generations, storages, converters and grids. As further explained in section 6.3, they can be added to the microgrid thanks to the "*add!()*" function.

```julia
mutable struct Microgrid
    parameters::GlobalParameters
    demands::Vector{AbstractDemand}
    generations::Vector{AbstractGeneration}
    storages::Vector{AbstractStorage}
    converters::Vector{AbstractConverter}
    grids::Vector{AbstractGrid}

    Microgrid(; parameters = GlobalParameters(8760, 20, 1)) = new(parameters)
end

```

Listing 6.1: Type definition for "Microgrid".

[1]In the toolbox, DES are called "Microgrid" without making the difference introduced in section 1.1.

- **Scenario**: the "Scenario" structure (in listing 6.2) is built from both the input time series and investment costs (stored in a ".jld" file), and the "Microgrid" component. A scenario is a sequence of realization of the uncertain parameters, associated with a set of investment costs for design decisions. The toolbox also provides a set of methods to manipulate these scenarios. For instance, the "*generate()*" and "*reduce()*" methods can be used to generate forecasts of the uncertain parameters and reduce the initial scenario set, respectively. Several approaches are given for both functionalities.

```
mutable struct Scenarios{T, O, I} <: AbstractScenarios
    demands::Vector{NamedTuple{(:t, :power),Tuple{T,O}}}
    generations::Vector{NamedTuple{(:t, :power, :cost), Tuple{T, O, I}}}
    storages::Vector{NamedTuple{(:cost,), Tuple{I}}}
    converters::Vector{NamedTuple{(:cost,), Tuple{I}}}
    grids::Vector{NamedTuple{(:cost_in, :cost_out), Tuple{O, O}}}
end

```

Listing 6.2: Type definition for "Scenario".

- **Designer**: the "Designer" structure is associated with the design strategy. A designer is initialized thanks to the "*initialize_designer!()*" method, which has the "Microgrid" and "Scenario" structures as input. More details are given in section 6.2.2.

- **Controller**: the "Controller" is associated with the operation strategy. As for the designer, a controller is initialized thanks to the "*initialize_controller!()*" method. More details are given in section 6.2.3.

Based on these components, the toolbox provides a set of methods to evaluate designer and controller performances, and to visualize the techno-economic indicators (see section 6.2.4 for more details). Figure 6.1 gives an overview of the Genesys.jl main functions based on the framework developed throughout this thesis.

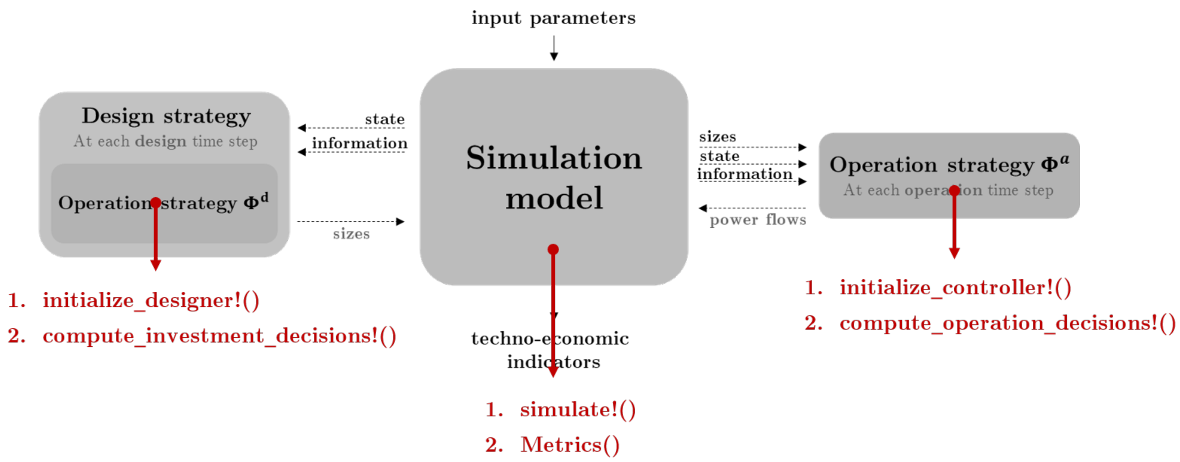


Figure 6.1: Overview of the main toolbox functions based on the framework developed throughout this thesis.

6.2.2 Implementation of a designer

The objective of this section is to describe the designer structure regarding the toolbox requirements.

Essential ingredients

A designer is defined as a mutable structure with at least a ”decisions” field and an inner constructor. The former is a named tuple with 3 entries (i.e., generations, storages and converters) where design decisions are assigned. Then, each designer comes with two essential functions:

1. *initialize_designer!()*: this function is needed to preallocate the ”decisions” field and derive all the ”offline” computations (e.g., calibration of the algorithm). Also, design decisions made in the first year are computed within this function. In order to be fully consistent with the methodological philosophy of this thesis, these first-year decisions should have been computed within the next function which corresponds to the design strategy in itself. The assumption behind the current implementation is that the simulation phase only begins once the size of the assets has been fixed in the first year.

2. *compute_investment_decisions!()*: this function is the design strategy that gives, at each design time step, the sizing decisions as a function of the state of the system and the available information. According to the previous remark, this function is only used when dealing with dynamic design issues where multiple decisions have to be made over the horizon.

To illustrate the above, the following parts describe the main design strategies implemented in this thesis. Because this work mainly addresses single-year problems, only the *initialize_designer!()* function is shown for each design method.

Manual designer

The manual designer is needed when the size of the assets is manually fixed by the user. The sizing values are given as input values when a designer object is built. The type definition is given by listing 6.3.

```
mutable struct Manual <: AbstractDesigner
    generations::Vector{Float64}
    storages::Vector{Float64}
    converters::Vector{Float64}
    decisions::NamedTuple

    Manual(; generations = [0.], storages = [0.], converters = [0.]) = new(generations, storages, converters)
end
```

Listing 6.3: Type definition of the manual designer

Then, the offline function (given by listing 6.4) basically preallocates the "decisions" field and assigns the design values previously filled in by the user (lines 5-14).

```
1 function initialize_designer!(mg::Microgrid, designer::Manual, w::AbstractScenarios)
2     # Preallocation
3     preallocate!(mg, designer)
4 
5     # Fix initial values
6     for (k, a) in enumerate(designer.decisions.generations)
7         a[1,:] .= designer.generations[k]
8     end
9     for (k, a) in enumerate(designer.decisions.storages)
10        a[1,:] .= designer.storages[k]
11    end
12    for (k, a) in enumerate(designer.decisions.converters)
13        a[1,:] .= designer.converters[k]
14    end
15
16    return designer
17 end
```

Listing 6.4: Initialize function of the manual designer

Mathematical programming designer

The mathematical programming designer (also called "MILP" in the toolbox) solves a single large optimization problem introduced in chapter 4. The type definition is given by listing 6.5.

```julia
mutable struct MILP <: AbstractDesigner
    options::MILPOptions
    decisions::NamedTuple
    model::JuMP.Model
    history::AbstractScenarios

    MILP(; options = MILPOptions()) = new(options)
end
```

Listing 6.5: Type definition of the mathematical programming designer

Besides the "decisions" field, the designer also has a "model" field where the JuMP model is stored, a "history" field where historical scenario can be stored for online computations and an "options" field where several options can be passed during initialization (e.g., solver type, scenario reducer, risk measures). The latter are given in a "MILPOptions" structure (with default field values already assigned) depicted in listing 6.6.

```julia
mutable struct MILPOptions
  solver::Module
  reducer::AbstractScenariosReducer
  objective_risk::AbstractRiskMeasure
  share_risk::AbstractRiskMeasure
  reopt::Bool
  read_reduction::Union{String, Nothing}
  write_reduction::Union{String, Nothing}

  MILPOptions(; solver = CPLEX,
              reducer = FeatureBasedReducer(),
              objective_risk = Expectation(),
              share_risk = Expectation(),
              reopt = false,
              read_reduction = nothing,
              write_reduction = nothing) =
              new(solver, reducer, objective_risk, share_risk, reopt, read_reduction, write_reduction)
end
```

Listing 6.6: Options of the mathematical programming designer

Then, the initialization function where the design values are computed is given in listing 6.7.

```julia
function initialize_designer!(mg::Microgrid, designer::MILP, w::Scenarios)
    # Preallocate
    preallocate!(mg, designer)

    # Scenario reduction from the optimization scenario pool
    if isa(designer.options.read_reduction, Nothing)
        println("Starting scenario reduction...")
        w_reduced, probabilities = reduce(designer.options.reducer, w)
        # Saving
        if !isa(designer.options.write_reduction, Nothing)
            save(designer.options.write_reduction, "scenarios", w_reduced, "probabilities", probabilities)
        end
    else
        println("Reading scenario reduction from file...")
```

```
15         w_reduced = load(designer.options.read_reduction, "scenarios")
16         probabilities = load(designer.options.read_reduction, "probabilities")
17     end
18 
19     # Initialize model
20     println("Building the model...")
21     designer.model = build_model(mg, designer, w_reduced, probabilities)
22 
23     # Compute investment decisions for the first year
24     println("Starting optimization...")
25     optimize!(designer.model)
26 
27     # Assign values
28     for k in 1:length(mg.generations)
29         designer.decisions.generations[k][1,:] .= value(designer.model[:r_g][k])
30     end
31     for k in 1:length(mg.storages)
32         designer.decisions.storages[k][1,:] .= value(designer.model[:r_sto][k])
33     end
34     for k in 1:length(mg.converters)
35         designer.decisions.converters[k][1,:] .= value(designer.model[:r_c][k])
36     end
37 
38     # Save history
39     designer.history = w_reduced
40 
41      return designer
42 end
```

Listing 6.7: Initialize function of the mathematical programming designer

We now comment on each line of the code:

Line 5-17: If the "read_reduction" option is filled in with a pathname, a reduced set of scenarios with the associated probabilities are loaded from the given path. Otherwise, the initial scenario set is reduced using the method specified in the designer options. The "write_reduction" option enables to save the reduced set.

Lines 19-25: The two-stage LP model is built using the JuMP library and solved by the solver specified in the options. Note that the JuMP model is automatically built from the microgrid architecture defined at the beginning of a study. The model function is given by listing 6.8. The reader could refer to the Github folder for more information about the functions associated with the model creation.

Lines 27-39: Design decisions are assigned to the "decisions" designer field and the reduced set of scenarios is saved for online computations.

```
function build_model(mg::Microgrid, designer::MILP, w::Scenarios, probabilities::Vector{Float64})
    # Sets
    nh, ns = size(w.demands[1].power, 1), size(w.demands[1].power, 3)
    # Initialize
    m = Model(designer.options.solver.Optimizer)
    # Add design decision variables
    add_investment_decisions!(m, mg.generations)
    add_investment_decisions!(m, mg.storages)
    add_investment_decisions!(m, mg.converters)
```

```
    # Add operation decision variables
    add_operation_decisions!(m, mg.storages, nh, ns)
    add_operation_decisions!(m, mg.converters, nh, ns)
    add_operation_decisions!(m, mg.grids, nh, ns)
    # Add design constraints
    add_investment_constraints!(m, mg.generations)
    add_investment_constraints!(m, mg.storages)
    add_investment_constraints!(m, mg.converters)
    # Add technical constraints
    add_technical_constraints!(m, mg.storages, mg.parameters.dh, nh, ns)
    add_technical_constraints!(m, mg.converters, nh, ns)
    add_technical_constraints!(m, mg.grids, nh, ns)
    # Add periodicity constraint
    add_periodicity_constraints!(m, mg.storages, ns)
    # Add power balance constraints
    add_power_balance!(m, mg, w, Electricity, nh, ns)
    add_power_balance!(m, mg, w, Heat, nh, ns)
    add_power_balance!(m, mg, w, Hydrogen, nh, ns)
    # Renewable share constraint
    add_renewable_share!(m, mg, w, probabilities, designer.options.share_risk, nh, ns)
    # Objective
    add_design_objective!(m, mg, w, probabilities, designer.options.objective_risk, nh, ns)
    return m
end
```

Listing 6.8: Model creation of the mathematical programming designer

Metaheuristic designer

The metaheuristic designer introduced in chapter 4 is given by listing 6.9. As previously, several options can be passed during initialization, and the reader could refer to the Github folder for in-depth details. Note that the ”results” field is an instance of ”MetaheuristicResults” where the ”Metaheuristics” library implements the niching algorithm introduced in chapter 4.

```
mutable struct Metaheuristic <: AbstractDesigner
    options::MetaheuristicOptions
    decisions::NamedTuple
    results::Metaheuristics.MetaheuristicResults
    history::AbstractScenarios

    Metaheuristic(; options = MetaheuristicOptions()) = new(options)
end
```

Listing 6.9: Type definition of the metaheuristic designer

Then, the initialization function is given by listing 6.10. As the method shares common lines with the mathematical programming method, only the differences are depicted in the listing. Hence, once the design bounds are fixed (line 5), the ”*optimize*” method of the ”Metaheuristic” library (`https://github.com/hradet/Metaheuristics.jl`) is called where the objective function (”fobj” in line 12) simulates the microgrid over the reduced set of scenarios, with the operation strategy specified in the designer options (see the Github folder for more details).

```
function initialize_designer!(mg::Microgrid, designer::Metaheuristic, w::Scenarios)
    ... ["Preallocation and scenario reduction"] ...

    # Bounds
    lb, ub = set_bounds(mg)

    # Optimize
    designer.results = Metaheuristics.optimize(lb, ub,
                                               designer.options.method,
                                               options = Metaheuristics.Options(
    iterations = designer.options.iterations, multithreads = designer.options.
    multithreads)
    ) do decisions
        fobj(decisions, mg, designer, w_reduced, probabilities)
      end

    ... ["Assignment of the design decisions"] ...

    return designer
end
```

Listing 6.10: Initialize function of the metaheuristic designer

6.2.3 Implementation of a controller

As in the previous section, the objective of this part is to describe the controller structure regarding the toolbox requirements.

Essential ingredients

A controller is defined as a mutable structure with at least a "decisions" field and an inner constructor. The former is a named tuple with 2 entries (i.e., storages and converters) where operation decisions are stored. Then, each controller comes with two essential functions:

1. *initialize_controller!()*: this function is needed to preallocate the "decisions" field and derive all the "offline" computations (e.g, model creation, initialize the scenario generator).

2. *compute_operation_decisions!()*: this function is the operation strategy which gives at each time step, the operation decisions as a function of the state of the system and the available information.

To illustrate the above, two controllers (i.e., RB and OLFC) are described in the following.

RB controller

Similar to the designer definition, the RB controller structure is given by listing 6.11. However, in this case, the "decisions" field is a named tuple with 3-dimensional arrays.

```
mutable struct RBC <: AbstractController
    options::RBCOptions
    decisions::NamedTuple
    history::AbstractScenarios

    RBC(; options = RBCOptions()) = new(options)
end
```

Listing 6.11: Type definition for the RB controller

The initialization function (listing 6.12) only preallocates the controller as no offline computation is required for the RB policy. In case some parameters need to be calibrated, this function can be used for that purpose.

```
function initialize_controller!(mg::Microgrid, controller::RBC, w::AbstractScenarios)
    # Preallocation
    preallocate!(mg, controller)

    return controller
end
```

Listing 6.12: Initialize function for the RB controller

In the toolbox, RB policies are specific to microgrid configurations. In practice, it means that each time a new microgrid topology is investigated, a new policy must be built by the user (unlike the OLFC strategy which automatically builds the appropriate JuMP model from the input microgrid architecture).

Three policies have been already implemented for the purpose of this work: the first one concerns the multi-energy system introduced in chapter 2, the second is for the toy-example which only considers electrical demand, and the third is used by the simplified energy system introduced in section 4.6. As the simulation progresses, operation decisions are computed at each time step by the operation strategy function described by listing 6.13. The corresponding policy is chosen according to the user option. Each "phi" function computes the operation decisions as a function of the state of the microgrid and available information. In case a new RB strategy needs to be built, the corresponding "if-else" rule has to be implemented in this latter function.

```
function compute_operation_decisions!(h::Int64, y::Int64, s::Int64, mg::Microgrid,
    controller::RBC)
    # Chose policy
    if controller.options.policy_selection == 1
        return phi_1(h, y, s, mg, controller)
    elseif controller.options.policy_selection == 2
        return phi_2(h, y, s, mg, controller)
    elseif controller.options.policy_selection == 3
        return phi_3(h, y, s, mg, controller)
```

```julia
    else
        println("Policy not defined !")
    end
end
```

Listing 6.13: RB policy

OLFC controller

The OLFC controller introduced in chapter 4, is given by listing 6.14. The OLFC parameters (i.e., scenario generator, number of scenarios, lookahead horizon, penalization coefficients, risk measure, reference trajectory) are set by the user in the OLFC options. Also, the size of the assets is assigned to the ”generations”, ”storages” and ”converters” fields (as for the manual designer), in order to build the corresponding LP model during the offline phase.

```julia
mutable struct OLFC <: AbstractController
    options::OLFCOptions
    generations::Vector{Float64}
    storages::Vector{Float64}
    converters::Vector{Float64}
    decisions::NamedTuple
    model::JuMP.Model
    history::AbstractScenarios
    OLFC(; options = OLFCOptions(),
           generations = [0.],
           storages = [0.],
           converters = [0.]) =
           new(options, generations, storages, converters)
end
```

Listing 6.14: Type definition for the OLFC controller

Then, the offline initialization function (given in listing 6.15) aims at building the model (line 6) and the scenario generator (lines 7-12), later used to forecast the uncertain parameters over the lookahead horizon.

```julia
1  function initialize_controller!(mg::Microgrid, controller::OLFC, w::Scenarios)
2      # Preallocate
3      preallocate!(mg, controller)
4      # Build model
5      println("Building the model...")
6      controller.model = build_model(mg, controller, w)
7      # Scenario reduction
8      println("Starting scenario reduction...")
9      w_reduced = reduce(controller.options.reducer, w)[1]
10     # Compute markov chain for scenario generation
11     println("Initializing scenario generator...")
12     controller.options.generator = initialize_generator!(controller.options.generator, [a
        for a in w_reduced.generations]..., [a for a in w_reduced.demands]...)
13     # History
14     controller.history = w
15     return controller
```

```
end
```

Listing 6.15: Initialize function of the OLFC controller

Similar to the mathematical programming designer, the model creation is depicted in listing 6.16 where design decisions are fixed variables (line 12). Note that only the power balances and grid powers depend on the number of scenarios ns, as previously explained in chapter 4.

```
1  function build_model(mg::Microgrid, controller::OLFC, w::Scenarios)
2      # Sets
3      nh, ns = controller.options.horizon, controller.options.nscenarios
4      # Initialize
5      m = Model(controller.options.solver.Optimizer)
6      set_optimizer_attribute(m,"CPX_PARAM_SCRIND", 0)
7      # Add investment variables
8      add_investment_decisions!(m, mg.generations)
9      add_investment_decisions!(m, mg.storages)
10     add_investment_decisions!(m, mg.converters)
11     # Fix their values
12     fix_investment_decisions!(m, controller.generations, controller.storages, controller.
       converters)
13     # Add operation decision variable with the non-anticipative structure
14     add_operation_decisions!(m, mg.storages, nh, 1)
15     add_operation_decisions!(m, mg.converters, nh, 1)
16     # Add demand and generation variables to be fixed online, along with recourse
       variable (grid)
17     add_operation_decisions!(m, mg.generations, nh, ns)
18     add_operation_decisions!(m, mg.demands, nh, ns)
19     add_operation_decisions!(m, mg.grids, nh, ns)
20     # Add technical constraints
21     add_technical_constraints!(m, mg.storages, mg.parameters.dh, nh, 1)
22     add_technical_constraints!(m, mg.converters, nh, 1)
23     add_technical_constraints!(m, mg.grids, nh, ns)
24     # Add power balance constraints
25     add_power_balance!(m, mg, w, Electricity, nh, ns, ispnet = true)
26     add_power_balance!(m, mg, w, Heat, nh, ns, ispnet = true)
27     add_power_balance!(m, mg, w, Hydrogen, nh, ns, ispnet = true)
28     return m
29 end
```

Listing 6.16: Model creation for the OLFC controller

Finally, the OLFC policy is given in listing 6.17:

Line 2-7: The uncertain parameters are forecasted over the lookahead horizon, and their values are fixed in the model previously built offline.

Line 8: The objective function is set with the risk measure defined by the user.

Line 11: The optimization problem is solved over the lookahead horizon.

Lines 12-24: Only the first value is kept and assigned to the corresponding "decisions" field.

```
1  function compute_operation_decisions!(h::Int64, y::Int64, s::Int64, mg::Microgrid,
       controller::OLFC)
2      # Forecast window
3      window = h:min(mg.parameters.nh, h + controller.options.horizon - 1)
```

```
    # Compute forecast
    demands, generations, grids, probabilities = compute_forecast(h, y, s, mg, controller, window)
    # Fix model variables
    fix_operation_variables!(controller, demands, generations, [a.soc[h,y,s] * a.Erated[y,s] for a in mg.storages])
    # Add objective
    add_objective!(controller, mg, grids, probabilities, window)
    # Optimize
    optimize!(controller.model)
    # Assign controller values
    for k in 1:length(mg.storages)
        controller.decisions.storages[k][h,y,s] = value(controller.model[:p_dch][1,1,k] - controller.model[:p_ch][1,1,k])
    end
    for (k,a) in enumerate(mg.converters)
        if a isa Heater
            controller.decisions.converters[k][h,y,s] = - value(controller.model[:p_c][1,1,k])
        elseif a isa Electrolyzer
            controller.decisions.converters[k][h,y,s] = - value(controller.model[:p_c][1,1,k])
        elseif a isa FuelCell
            controller.decisions.converters[k][h,y,s] = value(controller.model[:p_c][1,1,k])
        end
    end
end
```

Listing 6.17: OLFC policy

6.2.4 Simulation of a controller and designer

Once the designer and operation strategies have been built, the toolbox provides a function to evaluate these strategies by simulating the microgrid over multiple scenarios of a given horizon.

Algorithm 1: Simulation algorithm

Input : A microgrid, a controller, a designer, a scenario set Ω

Output: State and decision variables for both the design and operation for each hour h, year y and scenario s

```
for s = 1, ..., S do
    for y = 1, ..., Y do
        for h = 1, ..., H do
            → Update operation informations
            → Compute operation decisions
            → Compute operation dynamics
            → Compute power balances
        end
        → Update design informations
        → Compute design decisions
        → Compute design dynamics
    end
end
```

Note that the available information is updated at the beginning of each time step for both the operation and the design so that the controller and designer only have access to past and current observations. To this end, a difference is made between the values assigned to the ”Scenario” structure (i.e., all the values are available over the entire horizon) and the corresponding asset of the microgrid (where the values are updated time step after time step). For instance, both the ”Scenario” and ”Microgrid” structures have a ”Demand” field, however, the values of the microgrid ”Demand” are updated at each time step, from the ”Scenario” structure. Decisions are only made based on the ”Demand” attribute of the ”Microgrid” structure so that future values are unknown.

As the simulation progresses, the state and control values for both the design and the operation are assigned to the corresponding microgrid asset fields. Once the simulation is finished, the techno-economic metrics can be computed thanks to the *Metrics()*. The toolbox also provides a set of additional functions to visualize the main results.

6.3 Application on a case study

As an example, the two-stage model is solved for the toy-example (depicted in figure 6.2) using the mathematical programming design approach. Then, the sizing values are assessed with a predefined RB policy (corresponding to the second RB policy in the

Genesys.jl library) where the battery is charged as long as there is energy surplus and discharged otherwise.

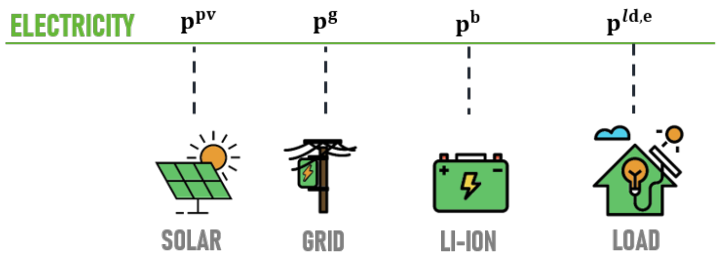


Figure 6.2: Schematic view of the toy-example.

Each line of listing 6.18 is commented in the following:

Line 2: Packages are loaded to run the application. "JLD" and "Dates" are used to load the input ".jld" file.

Lines 5-8: The hourly and yearly horizons (i.e., "nh" and "ny", respectively) and the number of scenarios "ns" are set as constant parameters. Also, the Ausgrid dataset (stored in a ".jld" file) is loaded into the namespace.

Lines 11-14: The microgrid is initialized with a renewable share constraint equal to 1. Then, the assets are added to the microgrid thanks to the *add!()* function. Note that both the "Grid" and "Demand" structures require to specify the type of the energy carrier.

Line 17: Two sets of scenarios for both the design and simulation are built from the input dataset and the microgrid structure.

Line 20: The designer is initialized using the mathematical programming method (called "MILP" in the library) with the default options.

Line 23: The controller is initialized using the second RB policy.

Line 26: The designer and controller are evaluated over the simulation scenario set which is different from the one used to design the microgrid. Note that the "multithreads" option is enabled to speed up computation.

Lines 29-35: Finally, techno-economic indicators are computed afterward and the main simulation results can be visualized.

```
1  # Load packages
2  using Genesys, JLD, Dates
3
4  # Parameters of the simulation
5  const nh, ny, ns = 8760, 2, 1000
6
7  # Load input data
8  data = load(joinpath("data","ausgrid_5_twostage.jld"))
9
10 # Initialize the microgrid
11 microgrid = Microgrid(parameters = GlobalParameters(nh, ny, ns, renewable_share = 1.))
```

```julia

# Add the equipment to the microgrid
add!(microgrid, Solar(), Grid(carrier = Electricity()), Liion(), Demand(carrier =
    Electricity()))

# Initialize scenarios
w_d, w_a = Scenarios(microgrid, data["w_d"]), Scenarios(microgrid, data["w_a"])

# Initialize the designer
designer = initialize_designer!(microgrid, MILP(), w_d)

# Initialize the controller
controller = initialize_controller!(microgrid, RBC(options = RBCOptions(policy_selection
    = 2)), w_d)

# Assessment
simulate!(microgrid, controller, designer, w_a, options = Genesys.Options(mode = "
    multithreads"))

# Compute the metrics
metrics = Metrics(microgrid, designer)

# Plot the metric statistics...
plot_metrics(metrics)

# ...and the operation!
plot_operation(microgrid)
```

Listing 6.18: Application on the toy example with the mathematical programming design method and a rule-based policy for the assessment.

6.4 How to use custom designer and controller methods within the framework?

The first step is to define your own designer and controller as Julia composite types (see listing 6.19). For this example, the designer is named ”bar” and the controller ”foo”. Do not forget to specify the subtype of each structure (i.e., ”AbstractController” and ”AbstractDesigner”) otherwise, they won't be recognized by the simulator. Moreover, both the ”decisions” structure field (defined as named tuple) and the inner constructor are mandatory for the two Julia types.

```julia
# Define your own designer...
mutable struct bar <: Genesys.AbstractDesigner
    decisions::NamedTuple
    bar() = new()
end

# ...and controller
mutable struct foo <: Genesys.AbstractController
    decisions::NamedTuple
    foo() = new()
```

```
end
```

Listing 6.19: Type definition for custom designer and controller.

Then, similarly to sections 6.2.2 and 6.2.3, you have to define the ”offline” methods which are used (at least!) to preallocate the ”decisions” field of each structure (listing 6.20). In this case, dummy controller and designer are implemented where the ”decisions” named tuple is filled with zeros.

```
# Define the "offline" functions for both the designer and controller
function Genesys.initialize_controller!(mg::Microgrid, controller::foo, w::Scenarios)
    # Preallocate
    Genesys.preallocate!(mg, controller)
    return controller
end

function Genesys.initialize_designer!(mg::Microgrid, designer::bar, w::Scenarios)
    # Preallocate
    Genesys.preallocate!(mg, designer)
    return designer
end
```

Listing 6.20: Offline functions for custom designer and controller.

Next, you have to implement the ”online” strategies for both the designer and controller, which give at each time step the decision variables as a function of the current system state and available information (listing 6.21). Again, the dummy controller and designer always return zeros regardless of the input information for this example.

```
# Define the "online" functions for both the designer and controller
function Genesys.compute_operation_decisions!(h::Int64, y::Int64, s::Int64, mg::Microgrid
    , controller::foo)
    return controller
end

function Genesys.compute_investment_decisions!(y::Int64, s::Int64, mg::Microgrid,
    designer::bar)
    return designer
end
```

Listing 6.21: Online functions for custom designer and controller.

Finally, you can evaluate the performance of the *foo()* controller and *bar()* designer by following the procedure previously described in section 6.3 (listing 6.22).

```
# Load packages
using Genesys, JLD, Dates

# Parameters of the simulation
const nh, ny, ns = 8760, 2, 1000

# Load input data
data = load(joinpath("data","ausgrid_5_twostage.jld"))

# Initialize the microgrid
microgrid = Microgrid(parameters = GlobalParameters(nh, ny, ns))
```

```julia

# Add the equipment to the microgrid
add!(microgrid, Solar(), Grid(carrier = Electricity()), Liion(), Demand(carrier =
    Electricity()))

# Initialize scenarios
w_d, w_a = Scenarios(microgrid, data["w_d"]), Scenarios(microgrid, data["w_a"])

# Initialize designer
designer = initialize_designer!(microgrid, bar(), w_d)

# Initialize controller
controller = initialize_controller!(microgrid, foo(), w_d)

# Assessment
simulate!(microgrid, controller, designer, w_a)
```

Listing 6.22: Application with the custom designer and controller.

6.5 Conclusion and perspectives

Genesys.jl has been introduced as a generic Julia toolbox to easily assess and compare different design and operation strategies for energy systems. First, the main toolbox ingredients were described, followed by the description of the design and operation strategies implemented in this thesis. Then, an example of how to use the toolbox was presented on a simple case study. Finally, the integration of custom designer and controller within the Genesys.jl framework was shown for anyone interested in this concern.

Although the general code structure is approximately settled, several critical points need further developments to improve the toolbox performance. A non-exhaustive enumeration is listed below (with increasing difficulty according to the author's point of view):

- Compute the first-year design decisions in the *compute_investment_decisions!()* function in order to be fully consistent with the methodological philosophy of the thesis. The *initialize_designer!()* function should only be dedicated for offline computations.

- Improve the toolbox input/output interfaces. Especially, 1) efficient scenario generation from ".csv" like files (which exchange format is the most appropriate?); 2) implementation of a module to save the outputs of any Genesys.jl study for later post-processing.

- Add more complex asset models to the library, which can be used in both the design and simulation phases.

- Add transmission line models for each carrier to account for energy transmission issues (cf the next bullet point).

- Add a spatial dimension in addition to the temporal and scenarios dimensions.
- Improve and take full advantage of parallel computing.
- Improve the simulator by interfacing the toolbox to more appropriate simulation tools (e.g., Modelica [139]) or even hardware-in-the-loop (HIL) energy systems.

Despite all these aforementioned limitations, the Genesys.jl toolbox can be a valuable ”starter pack” for researchers looking for practical implementation of several design and operation strategies for energy systems.

Chapter 7

Conclusion and perspectives

Contents

7.1 Summary

In recent years, driven by ambitious decarbonization plans, the energy system has been rapidly evolving to integrate a growing share of renewable energies. The shift from dispatchable to variable resources has led to the development of new flexibility means to overcome the mismatch between production and demand at all times. Among the different options, storage systems along with multi-energy strategies tend to be promising directions to mitigate the production variability by coupling the energy carriers with each other.

As introduced in chapter 1, planning the design of such systems is a challenging task because the problem displays multiple facets that are difficult for policy- and decision-makers to address in a systemic manner. Also, decisions are made while many parameters remain uncertain as their values progressively unfold over time. Therefore, mathematical tools are often needed to provide decision support regarding several techno-economic requirements: the problem is expressed in a form of optimization problems where decision variables are the sizes of the equipment. The resulting problem is most of the time intractable, thus simplifications are inevitable to come up with solutions. The main modeling challenge is then to determine the relevant simplifications regarding the objective of the study.

Within this context, chapter 2 developed a simple deterministic model to assess the value of multi-energy systems and seasonal storage to supply residential customers with a high share of solar production. The objective of this part was threefold: i) such optimization problems are easily solved with standard solvers so that fast parametric studies can be run to identify global DES design and techno-economic trends; ii) the result of the deterministic problem with perfect foresight gives the best (but unreachable) solution for a given scenario, which is valuable information to understand the topology of the problem; iii) the deterministic model forms the basis of the stochastic formulation, thus this step is an essential prerequisite to properly understand the next chapter developments. Results of the parametric analysis showed that the total annual cost increases exponentially with the renewable ratio and the integration of a low-cost dispatchable source may significantly alleviate the overall system cost. Furthermore, hydrogen units only emerge for high shares of solar production (i.e., renewable share greater than 80%). In the particular case of standalone DES, the multi-energy system with seasonal storage leads to better techno-economic performance than the battery-only solution. In this case, cogenerated heat from hydrogen helps reduce the total cost of the system.

Then, the next two chapters dealt with the design of the DES under uncertainties: the energy demands, the solar production and the electricity tariff are not perfectly known when sizing the system. The first objective was to develop a method to generate synthetic scenarios of energy demands and production for both long and short-term applications. To this end, a simple stochastic model based on Markov chains was presented in chapter 3. The method was applied to a residential case study and the results showed good performance to recover the main statistical features of the initial dataset while introducing temporal variability between scenarios. The drawback of such a method is that the synthetic future scenarios only display power levels and daily patterns from the initial dataset. However, in practice, the epistemic nature of future scenarios may be different from that of the historical dataset. Despite this limitation, the method is appropriate for the modeling purpose of this thesis where a large set of scenarios is needed to assess the performance of different design and operation methodologies.

Next, chapter 4 took on the challenge of developing several design (i.e., mathematical programming and metaheuristic approaches in a two-stage fashion) and operation strategies (i.e., anticipative, rule-based and open-loop feedback control) under uncertainties for a multi-energy system with seasonal storage. The main objective of this part was to challenge the perfect foresight hypothesis attached to typical mathematical programming approaches from the literature, regarding realistic operation strategies which only have access to past and current information (i.e., the future energy demands, production and electricity tariffs are not perfectly known when sizing the systems). Furthermore, this section examined more precisely the interaction between design and operation on a simplified case study. Results showed that the real-life operation strategy and the one embedded in the design procedure do not have to be strictly identical. However, their level of optimality must be similar, and this work helped quantify this notion. As a result, the mathematical programming approach (widely implemented in the literature to solve complex planning problems) is relevant if, and only if, the DES is finally operated with a real-time policy that performs similarly to the anticipative operation strategy. This latter condition might be critical for complex energy systems where the development of ”near-optimal” real-time policies is difficult. In the case of poor performance operation strategies, the sizing values resulting from the two-stage LP model might not be appropriate to meet the project requirements. In this case, the metaheuristic approach should be considered provided that the operation policy is fast enough to be integrated into the design loop. The results also demonstrated the relevance of stochastic solutions over deterministic designs as the latter can lead to misleading results since they strongly depend on the scenario selected for optimization.

Finally, the last part of this thesis (see chapter 5) addressed multi-year planning issues where multiple design decisions can be made over the horizon of the project. More precisely, a multi-time scale model (derived from the literature) was implemented on a simple case study (i.e., battery and solar panels) where the impact of the operation over equipment lifetimes is taken into account. Unlike the majority of studies from the literature, system lifetimes are not fixed *a priori*, but they depend on the way systems are operated over time. This aspect is particularly interesting for storage systems such as Li-ion batteries where the number of charge/discharge cycles has great consequences on their aging. This dynamic aware aging design method was compared to two heuristic design strategies based on the two-stage model previously introduced in chapter 4. Results showed that, unlike heuristic strategies, the dynamic aware aging method can control both the investment pathway and the battery aging, thereby reducing the overall cost of the DES.

All the previous works would not have been possible without a strong numerical background. Therefore, Genesys.jl was introduced in chapter 6 as a generic Julia toolbox to easily assess and compare different design and operation strategies for energy systems. The toolbox allows switching from one strategy to another without changing the overall code. The last part of this chapter presented how custom design and operation approaches can be easily added to the framework for anyone interested in such issues.

7.2 Contributions

Based on the previous developments, the contributions of this thesis are mainly methodological. They can be summarized as follows:

- **The introduction of a modeling framework to assess different design and operation strategies under uncertainty**. To this end, the models used to design and simulate the strategies are clearly divided into two distinct parts: the simulator is assumed to be a good representation of the real system, while the design model can be built on simplified versions of the problem. Also, scenarios used to design the DES and to compute the techno-economic indicators are different to avoid any bias in the evaluation.

- **The joint comparison of diverse design and realistic operation strategies under uncertainty**. As a result, the scope of the mathematical programming approach was clarified regarding the performance of realistic operation strategies. This work helped quantify this notion. Another method based on a metaheuristic algorithm was also provided to design DES that are controlled with poor performance policies.

- **The comparison of a dynamic aware aging design method with two heuristic design strategies based on a two-stage model**.

- **The introduction of a generic toolbox to easily assess and compare different design and operation strategies without changing the overall code**.

We don't have the arrogance to claim major advances through this work, but we hope that our little steps will contribute to the research effort on these questions.

7.3 Future directions

The work presented in this thesis comes with multiple limitations attached to the modeling choices. They were previously discussed in the introduction. In addition, this section attempts to identify some perspectives that seem crucial to the author in order to enhance the value of this work.

The first direction pertains to the simulator model. In this work, the same model granularity (e.g., technological details, time steps) was kept for both the design and simulator for the sake of clarity (and lack of time...). The only differences were deliberately introduced to highlight some specific research points. However, the simulator could be improved to come closer to the real behavior of the system by using appropriate simulation tools such as Modelica [139] or TRNSYS [87] (the simulator could also be replaced by hardware-in-the-loop (HIL) simulation). Then, the model implemented in the design strategy can be gradually simplified to evaluate and compare the impacts of the different simplifications over the results (i.e., both the resulting sizes and the techno-economic indicators computed by simulation).

Next, one obvious perspective would be to integrate both operation and investment uncertainties in the multi-year planning model. Also, the benefit of including the aging dynamic into the model must be discussed in view of typical multi-stage approaches where the asset lifetimes are fixed *a priori*. This latter assumption may be sufficient from a design perspective, and the aging dynamic may be unnecessarily complex for this purpose. Furthermore, this thesis only addressed uncertainties attached to time series. Another source of uncertainty comes from the nature of the modeling process. Indeed, a model is an approximation of reality, so any model is by nature uncertain. Integrating uncertainties related to the physical model of the assets could be a direction for future work.

Another direction that is related to the first point would be to compare "soft-linking" methodologies to the design methods developed in this thesis. Indeed, some simulation models already exist in the literature and it would be interesting to couple these models to design strategies where a parameter is progressively tuned to reach the requirements. The performance of these methods could then be compared to traditional design strategies, such as those developed in this work.

Based on the previous methodological improvements, it would be interesting to apply these methodologies to more complex case studies where DES users participate in multiple energy markets (e.g., frequency ancillary services, capacity mechanism, balancing mechanism) that are available to support VRE integration into the system. Future work could include performing techno-economic analyses in this context to determine new business models for these systems. And last, but not least, it would be crucial to include social and environmental indicators into the framework (e.g., based on life-cycle analysis) to assess the true relevance of such multi-energy systems to achieve net carbon reduction goals in accordance with societal needs.

The base of the pyramid has been built during this thesis. The edifice is huge, and the rest of the construction will be the role of future doctoral and postdoctoral students based on their own research issues.

Appendix A

The conditional value at risk (CVaR) risk measure

The CVaR is a coherent risk measure introduced by Rockafellar and Uryasev [140] depicted in figure A.1.

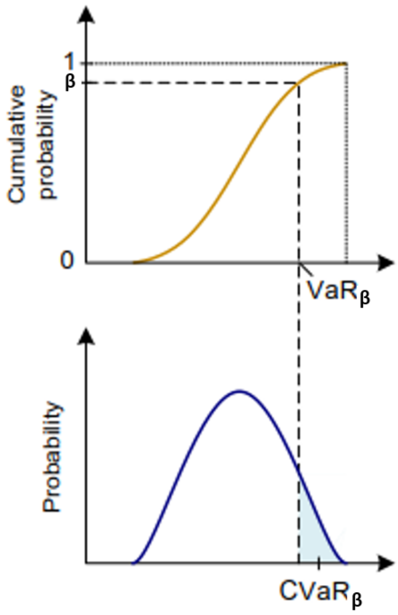


Figure A.1: Schematic representation of the VaR and CVaR of a random variable (from Mavromatidis et al [5]): $\beta = 0$ corresponds to the expectation while $\beta = 1$ is the worst case.

Given a random variable X and its associated probability distribution, the value at risk (VaR_β) at level $\beta \in [0, 1]$ is the β-quantile of X. Therefore, the CVaR_β at level β is defined as the expected values of X that are larger than the VaR:

$$\mathrm{CVaR}_\beta(X) = \mathbb{E}[X | X \geq \mathrm{VaR}_\beta(X)] \tag{A.1}$$

Basically, the CVaR is the expected value of the $(1 - \beta) \cdot 100\%$ worst scenarios. The advantage of CVaR is that it can easily be linearized to be integrated into optimization problems. This usually needs the introduction of auxiliary variables and the authors in [112] rigorously demonstrated the linearization procedure. Readers seeking more information about this point should directly refer to the reference.

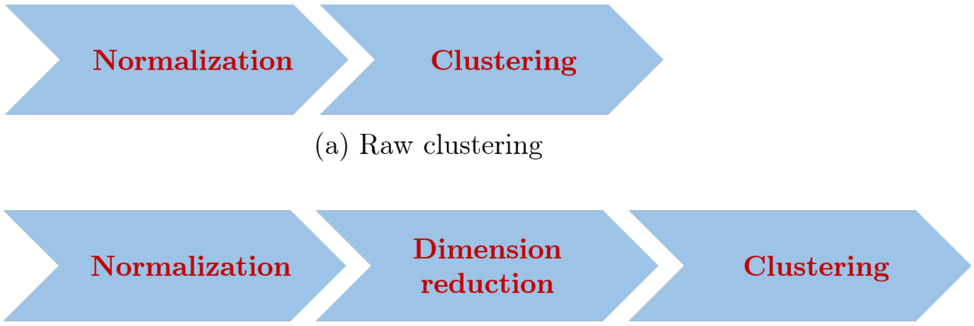
Appendix B

Scenario generation for stochastic programming

The main challenge in solving stochastic programs is to determine the right number of scenarios to represent uncertainties: the latter must be low enough to keep computational tractability while ensuring a good representation of the underlying stochastic processes.

Therefore, *scenario generation* methods can be classified into bottom-up and top-down approaches. In the former case, the goal is to build the sample space starting from a single scenario and then iteratively add scenarios until the model objective converges [75, 141]. On the other hand, top-down approaches start with a large set of scenarios, and the objective is to statistically approximate the initial set with a lower number of scenarios. In this case, heuristic strategies can be used to minimize a probability distance from the original distribution (e.g., forward selection and backward reduction) [107, 108, 142]. The number of scenarios can be fixed using stability tests [75]: the model is run with a growing number of scenarios extracted from one of the previous approaches until the objective function converges. According to King et al [75], **it is usually difficult to observe good convergence for large scale stochastic problems, especially over the solutions (i.e., sizing values in this work), because stochastic problems mostly have flat objectives where multiple solutions could lead to similar results**. Note that scenario generation methods could be a research topic in itself. The aim of this work is to provide two simple top-down heuristic algorithms to reduce the initial set of scenarios. The number of scenarios is then fixed through stability tests.

Therefore, the initial set of scenarios Ω^d is made of 1000 annual time series (at an hourly time step) of energy demands, production and electricity tariff. The two heuristic methods are introduced in the following and depicted in figure B.1.



(a) Raw clustering

(b) Feature based clustering

Figure B.1: The different reduction steps for the two heuristic strategies: (a) the ”raw clustering” approach directly applies the k-medoids algorithm on the initial set of scenarios while (b) the clustering algorithm is run over a limited set of features (i.e., sum, maximum and first 4th statistical moments) with the ”feature-based clustering” method.

The first step in the reduction procedure is to aggregate and normalize the data to keep the synchronicity between time series. Then, the first method implemented in this work (called ”raw clustering”) directly applies the clustering algorithm on the 4 (electricity, heat, PV and tariff) x 8760 (hours) x 1000 (scenarios) observations. Following the *data science* vocabulary, the initial set has 35 040 dimensions and 1000 observations. As it has been noticed several times in the literature [74, 115], applying clustering algorithms to such high-dimensional data may lead to inaccurate results because the concept of ”distance” becomes less precise. Therefore, the second method (called ”feature-based clustering”) integrates a typical intermediate step where the number of dimensions is reduced before applying the clustering method. The main challenge is then to determine the representative features that will constitute the low-dimensional space. These features must be chosen according to the purpose of the study (i.e., the design of DES in this work).

A large number of methods are available for both dimension reduction (e.g., the representative features can whether be chosen manually according to the intuition of the user or extracted using more sophisticated dimension reduction algorithms as PCA [143] or UMAP [144]) and clustering (e.g., k-medoids [92], DBSCAN [145]). In this work, both scenario reduction methods apply the k-medoids algorithm with the euclidean distance [1] to cluster the scenarios. However, with the second method, the dimension is first reduced by manually extracting a limited set of features for each time series (i.e., sum, maximum and first 4th statistical moments) following [74] and [114]. The probability of each scenario is computed based on the number of cluster assignments. Note that the more sophisticated techniques previously introduced as examples were also used but they did not provide better results.

[1]Dynamic Time Warping (DTW) might be more appropriate for time series but this latter was not studied in this work. The reader could refer to [146] for more information about this metric.

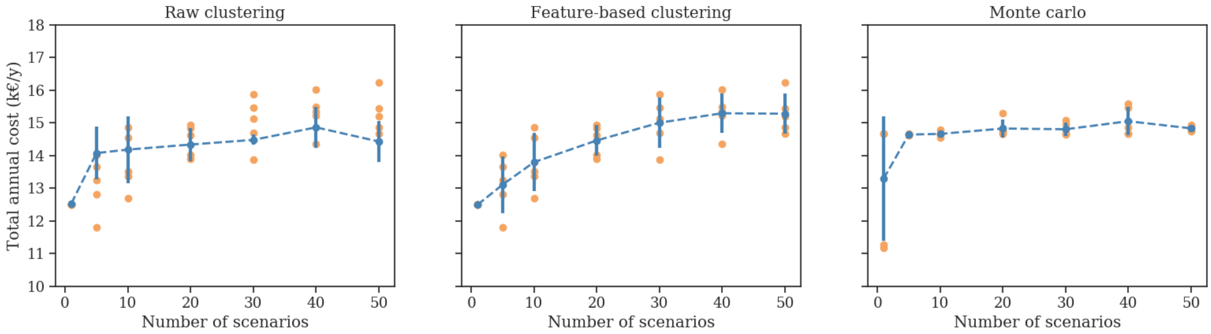


Figure B.2: Results of the stability tests for the two scenario reduction methods. The results are compared with those of the "Monte Carlo" approach where the scenarios are randomly selected from the initial dataset.

Figure B.2 shows the stability test results in a case where the risk measure is the expectation and the renewable ratio is equal to 1. To this end, the mathematical programming approach is run with a growing number of scenarios from both reduction methods. As these methods are based on aleatory numbers, they are run 10 times to deal with their stochastic nature. For each "number of scenarios", the corresponding error bar (i.e., mean and standard deviation) is drawn in the figure.

As with the authors' conclusions in [75], convergence is not clearly depicted in the figure with both methods. Even if the "feature-based" approach seems to reach a plateau at 40 scenarios, the variability between runs still remains around 10% of the mean value. The convergence for the "raw clustering" method is even more questionable and the gain of one reduction method over the other is not obvious. More surprisingly is the result of the Monte Carlo method. Indeed, while the value of the total annual cost is highly variable for a single scenario (see the error bar - this is consistent with the conclusion of section 4.5.3), the results seem to be more stable than for the other approaches when the number of scenarios increases. The main drawback of such a method is that the scenarios are randomly selected with equal probabilities. Therefore, there is a chance that the scenarios extracted by the Monte Carlo approach are not statistically representative of the initial dataset. This remark is even more relevant for a different risk measure that would be more sensitive to the tail of the distribution (e.g., the CVaR at 95%).

Despite these poor results, the feature-based method is used to constitute a reduced set of 40 scenarios as a trade-off between computation times and statistical representation. This method is chosen among the others as it seems to the author more accurate to approximate the entire probability distribution. Note that in any case, the consistency of the reduction approach will be settled in section 4.5 when the design values will be assessed on the out-of-sample simulator: if the techno-economic requirements are not met, the reduction method will have to be revised.

In order to be fully accurate, stability tests should have been run for each case study (i.e., risk measure and renewable share value) and for each design method (i.e., mathematical programming and metaheuristic). For computational reasons (each run takes about 2 hours when the number of scenarios is greater than 20...), the reduced scenario set is only built once with the mathematical programming approach and the same scenarios are used to run the parametric analyses of section 4.5.

Bibliography

[1] Valérie Masson-Delmotte et al., eds. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2021.

[2] IEA. *Net Zero by 2050 - A Roadmap for the Global Energy Sector*. en. Tech. rep. PAris, 2021, p. 224.

[3] IEA. *World Energy Outlook 2021*. en. Tech. rep. Paris, 2021, p. 386.

[4] RTE. *Futurs énergétiques 2050 : les scénarios de mix de production à l'étude permettant d'atteindre la neutralité carbone à l'horizon 2050*. fr. 2021.

[5] Georgios Mavromatidis. "Model-based design of distributed urban energy systems under uncertainty". en. Accepted: 2017-09-11T11:11:22Z. Doctoral Thesis. ETH Zurich, 2017. DOI: `10.3929/ethz-b-000182697`.

[6] Kari Alanne and Arto Saari. "Distributed energy generation and sustainable development". en. In: *Renewable and Sustainable Energy Reviews* 10.6 (Dec. 2006), pp. 539–558. ISSN: 1364-0321. DOI: `10.1016/j.rser.2004.11.004`.

[7] Pierluigi Mancarella. "MES (multi-energy systems): An overview of concepts and evaluation models". en. In: *Energy* 65 (Feb. 2014), pp. 1–17. ISSN: 0360-5442. DOI: `10.1016/j.energy.2013.10.041`.

[8] Peter D. Lund et al. "Review of energy system flexibility measures to enable high levels of variable renewable electricity". In: *Renewable and Sustainable Energy Reviews* 45 (May 2015), pp. 785–807. ISSN: 1364-0321. DOI: `10.1016/j.rser.2015.01.057`.

[9] IEA. *Distributed energy resources for net zero: An asset or a hassle to the electricity grid? – Analysis*. en-GB. Tech. rep. Paris, 2021.

[10] Ausgrid. *Solar home electricity data - Ausgrid*. en. 2020.

[11] Elizabeth L. Ratnam et al. "Residential load and rooftop PV generation: an Australian distribution network dataset". In: *International Journal of Sustainable Energy* 36.8 (Sept. 2017), pp. 787–806. ISSN: 1478-6451. DOI: `10.1080/14786451.2015.1100196`.

[12] D. Connolly et al. "A review of computer tools for analysing the integration of renewable energy into various energy systems". In: *Applied Energy* 87.4 (Apr. 2010), pp. 1059–1082. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2009.09.026.

[13] Miguel Chang et al. "Trends in tools and approaches for modelling the energy transition". en. In: *Applied Energy* 290 (May 2021), p. 116731. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2021.116731.

[14] Etienne Cuisinier et al. "Techno-economic planning of local energy systems through optimization models: a survey of current methods". en. In: *International Journal of Energy Research* n/a.n/a (2021). ISSN: 1099-114X. DOI: https://doi.org/10.1002/er.6208.

[15] Niina Helistö, Juha Kiviluoma, and Hannu Reittu. "Selection of representative slices for generation expansion planning using regular decomposition". en. In: *Energy* (Aug. 2020), p. 118585. ISSN: 0360-5442. DOI: 10.1016/j.energy.2020.118585.

[16] Georgios Mavromatidis et al. "Ten questions concerning modeling of distributed multi-energy systems". en. In: *Building and Environment* 165 (Nov. 2019), p. 106372. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2019.106372.

[17] Niina Helistö et al. "Including operational aspects in the planning of power systems with large amounts of variable generation: A review of modeling approaches". en. In: *WIREs Energy and Environment* 8.5 (2019), e341. ISSN: 2041-840X. DOI: 10.1002/wene.341.

[18] Warren B. Powell. "Clearing the Jungle of Stochastic Optimization". In: *Bridging Data and Decisions*. INFORMS TutORials in Operations Research. INFORMS, Sept. 2014, pp. 109–137. ISBN: 978-0-9843378-5-9. DOI: 10.1287/educ.2014.0128.

[19] Tristan Rigaut. "Time decomposition methods for optimal management of energy storage under stochasticity". en. PhD thesis. Université Paris-Est, May 2019.

[20] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. "Uncertainty and global sensitivity analysis for the optimal design of distributed energy systems". en. In: *Applied Energy* 214 (Mar. 2018), pp. 219–238. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.01.062.

[21] Mohammad Ameri and Zahed Besharati. "Optimal design and operation of district heating and cooling networks with CCHP systems in a residential complex". en. In: *Energy and Buildings* 110 (Jan. 2016), pp. 135–148. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.10.050.

[22] Stefano Bracco, Gabriele Dentici, and Silvia Siri. "DESOD: a mathematical programming tool to optimally design a distributed energy system". en. In: *Energy* 100 (Apr. 2016), pp. 298–309. ISSN: 0360-5442. DOI: 10.1016/j.energy.2016.01.050.

[23] Paolo Gabrielli et al. “Optimal design of multi-energy systems with seasonal storage”. en. In: *Applied Energy* 219 (June 2018), pp. 408–424. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.07.142.

[24] Bei Li et al. “Sizing of a stand-alone microgrid considering electric power, cooling/heating, hydrogen loads and hydrogen storage degradation”. In: *Applied Energy* 205 (Nov. 2017), pp. 1244–1259. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.08.142.

[25] H. K. Fathy et al. “On the coupling between the plant and controller optimization problems”. In: *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148).* Vol. 3. June 2001, 1864–1869 vol.3. DOI: 10.1109/ACC.2001.946008.

[26] Anya Heider et al. “Flexibility options and their representation in open energy modelling tools”. en. In: *Energy Strategy Reviews* 38 (Nov. 2021), p. 100737. ISSN: 2211-467X. DOI: 10.1016/j.esr.2021.100737.

[27] Kris Poncelet et al. “Impact of the level of temporal and operational detail in energy-system planning models”. en. In: *Applied Energy* 162 (Jan. 2016), pp. 631–643. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2015.10.100.

[28] Holger Teichgraeber and Adam R. Brandt. “Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison”. en. In: *Applied Energy* 239 (Apr. 2019), pp. 1283–1293. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.02.012.

[29] Stefan Pfenninger. “Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability”. en. In: *Applied Energy* 197 (July 2017), pp. 1–13. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.03.051.

[30] Oriol Raventós and Julian Bartels. “Evaluation of Temporal Complexity Reduction Techniques Applied to Storage Expansion Planning in Power System Models”. en. In: *Energies* 13.4 (Jan. 2020), p. 988. DOI: 10.3390/en13040988.

[31] Leander Kotzur et al. “Time series aggregation for energy system design: Modeling seasonal storage”. en. In: *Applied Energy* 213 (Mar. 2018), pp. 123–135. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.01.023.

[32] Nikolaos E. Koltsaklis and Athanasios S. Dagoumas. “State-of-the-art generation expansion planning: A review”. In: *Applied Energy* 230 (Nov. 2018), pp. 563–589. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.08.087.

[33] T. Brown et al. “Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system”. en. In: *Energy* 160 (Oct. 2018), pp. 720–739. ISSN: 0360-5442. DOI: 10.1016/j.energy.2018.06.222.

[34] Antonio J. Conejo et al. *Investment in Electricity Generation and Transmission*. en. Cham: Springer International Publishing, 2016. ISBN: 978-3-319-29499-5 978-3-319-29501-5. DOI: 10.1007/978-3-319-29501-5.

[35] Amélia Nadal. "Influence des incertitudes sur l'optimisation technico-économique de systèmes énergétiques hybrides". fr. PhD thesis. Communauté Université Grenoble Alpes, Nov. 2019.

[36] Martin Geidl. "Integrated Modeling and Optimization of Multi-Carrier Energy Systems". en. Doctoral Thesis. ETH Zurich, 2007. DOI: 10.3929/ethz-a-005377890.

[37] Hadi Sadeghi et al. "The energy hub: An extensive survey on the state-of-the-art". en. In: *Applied Thermal Engineering* 161 (Oct. 2019), p. 114071. ISSN: 1359-4311. DOI: 10.1016/j.applthermaleng.2019.114071.

[38] R. Rigo-Mariani et al. "A Combined Cycle Gas Turbine Model for Heat and Power Dispatch Subject to Grid Constraints". In: *IEEE Transactions on Sustainable Energy* (2019), pp. 1–1. ISSN: 1949-3029. DOI: 10.1109/TSTE.2019.2894793.

[39] Paolo Gabrielli, Matteo Gazzani, and Marco Mazzotti. "Electrochemical conversion technologies for optimal design of decentralized multi-energy systems: Modeling framework and technology assessment". en. In: *Applied Energy* 221 (July 2018), pp. 557–575. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.03.149.

[40] Jonas Allegrini et al. "A review of modelling approaches and tools for the simulation of district-scale energy systems". In: *Renewable and Sustainable Energy Reviews* 52 (Dec. 2015), pp. 1391–1404. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.07.123.

[41] Juan M. Morales et al. *Integrating Renewables in Electricity Markets: Operational Problems*. en. International Series in Operations Research & Management Science. Springer US, 2014. ISBN: 978-1-4614-9410-2.

[42] L. Gacitua et al. "A comprehensive review on expansion planning: Models and tools for energy policy analysis". In: *Renewable and Sustainable Energy Reviews* 98 (Dec. 2018), pp. 346–360. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.08.043.

[43] Carlos Gamarra and Josep M. Guerrero. "Computational optimization techniques applied to microgrids planning: A review". In: *Renewable and Sustainable Energy Reviews* 48 (Aug. 2015), pp. 413–424. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.04.025.

[44] Nestor A. (Sepulveda Morales) Sepulveda. "Decarbonization of Power Systems, Multi-Stage Decision-Making with Policy and Technology Uncertainty". eng. Thesis. Massachusetts Institute of Technology, 2020.

[45] Yixian Liu, Ramteen Sioshansi, and Antonio J. Conejo. “Multistage Stochastic Investment Planning With Multiscale Representation of Uncertainties and Decisions”. en. In: *IEEE Transactions on Power Systems* 33.1 (Jan. 2018), pp. 781–791. ISSN: 0885-8950, 1558-0679. DOI: 10.1109/TPWRS.2017.2694612.

[46] Emilio L. Cano, Javier M. Moguerza, and Antonio Alonso-Ayuso. “A multi-stage stochastic optimization model for energy systems planning and risk management”. en. In: *Energy and Buildings* 110 (Jan. 2016), pp. 49–56. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.10.020.

[47] Karl-Kiên Cao et al. “Classification and Evaluation of Concepts for Improving the Performance of Applied Energy System Optimization Models”. en. In: *Energies* 12.24 (Jan. 2019), p. 4656. DOI: 10.3390/en12244656.

[48] Stefan Pfenninger, Adam Hawkes, and James Keirstead. “Energy systems modeling for twenty-first century energy challenges”. en. In: *Renewable and Sustainable Energy Reviews* 33 (May 2014), pp. 74–86. ISSN: 1364-0321. DOI: 10.1016/j.rser.2014.02.003.

[49] Kristopher A. Pruitt, Robert J. Braun, and Alexandra M. Newman. “Evaluating shortfalls in mixed-integer programming approaches for the optimal design and dispatch of distributed generation systems”. en. In: *Applied Energy*. Special Issue on Advances in sustainable biofuel production and use - XIX International Symposium on Alcohol Fuels - ISAF 102 (Feb. 2013), pp. 386–398. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2012.07.030.

[50] Joseph DeCarolis et al. “Formalizing best practice for energy system optimization modelling”. en. In: *Applied Energy* 194 (May 2017), pp. 184–198. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.03.001.

[51] Paul Stadler, Araz Ashouri, and François Maréchal. “Model-based optimization of distributed and renewable energy systems in buildings”. en. In: *Energy and Buildings* 120 (May 2016), pp. 103–113. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2016.03.051.

[52] Yun Yang, Shijie Zhang, and Yunhan Xiao. “An MILP (mixed integer linear programming) model for optimal design of district-scale distributed energy resource systems”. en. In: *Energy* 90 (Oct. 2015), pp. 1901–1915. ISSN: 0360-5442. DOI: 10.1016/j.energy.2015.07.013.

[53] Portia Murray et al. “A comparison of storage systems in neighbourhood decentralized energy system applications from 2015 to 2050”. en. In: *Applied Energy* 231 (Dec. 2018), pp. 1285–1306. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.08.106.

[54] Amélia Nadal et al. “Uncertainty sensitivity assessment on the optimization of the design of complex energy systems: two complementary approaches”. en. In: (2019), p. 6.

[55] Stefano Moret et al. “Characterization of input uncertainties in strategic energy planning models”. In: *Applied Energy* 202 (Sept. 2017), pp. 597–617. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.05.106.

[56] Claudia D’Ambrosio, Andrea Lodi, and Silvano Martello. “Piecewise linear approximation of functions of two variables in MILP models”. en. In: *Operations Research Letters* 38.1 (Jan. 2010), pp. 39–46. ISSN: 01676377. DOI: 10.1016/j.orl.2009.09.005.

[57] J. F. Benders. “Partitioning procedures for solving mixed-variables programming problems”. en. In: *Numerische Mathematik* 4.1 (Dec. 1962), pp. 238–252. ISSN: 0945-3245. DOI: 10.1007/BF01386316.

[58] John Forrest et al. *coin-or/Cbc: Version 2.9.9*. July 2018. DOI: 10.5281/zenodo.1317566.

[59] IBM ILOG Cplex. “V12. 1: User’s Manual for CPLEX”. en-US. In: *International Business Machines Corporation* 46.53 (2009), p. 157.

[60] Mohammad H. Moradi, Mohsen Eskandari, and Hemen Showkati. “A hybrid method for simultaneous optimization of DG capacity and operational strategy in microgrids utilizing renewable energy resources”. en. In: *International Journal of Electrical Power & Energy Systems* 56 (Mar. 2014), pp. 241–258. ISSN: 0142-0615. DOI: 10.1016/j.ijepes.2013.11.012.

[61] A. Petrowski. “A clearing procedure as a niching method for genetic algorithms”. In: *Proceedings of IEEE International Conference on Evolutionary Computation*. May 1996, pp. 798–803. DOI: 10.1109/ICEC.1996.542703.

[62] Benjamin Guinot et al. “Techno-economic study of a PV-hydrogen-battery hybrid system for off-grid power supply: Impact of performances’ ageing on optimal system sizing and competitiveness”. en. In: *International Journal of Hydrogen Energy* 40.1 (Jan. 2015), pp. 623–632. ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2014.11.007.

[63] Jiangjiang Wang, Jun Sui, and Hongguang Jin. “An improved operation strategy of combined cooling heating and power system following electrical load”. en. In: *Energy* 85 (June 2015), pp. 654–666. ISSN: 0360-5442. DOI: 10.1016/j.energy.2015.04.003.

[64] Jun Zhang et al. “A bi-level program for the planning of an islanded microgrid including CAES”. In: *2015 IEEE Industry Applications Society Annual Meeting*. Oct. 2015, pp. 1–8. DOI: 10.1109/IAS.2015.7356783.

[65] J. P. Deane et al. "Soft-linking of a power systems model to an energy systems model". en. In: *Energy*. 8th World Energy System Conference, WESC 2010 42.1 (June 2012), pp. 303–312. ISSN: 0360-5442. DOI: 10.1016/j.energy.2012.03.052.

[66] Dominik Franjo Dominković et al. "Implementing flexibility into energy planning models: Soft-linking of a high-level energy planning model and a short-term operational model". en. In: *Applied Energy* 260 (Feb. 2020), p. 114292. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.114292.

[67] Matija Pavičević et al. "The potential of sector coupling in future European energy systems: Soft linking between the Dispa-SET and JRC-EU-TIMES models". en. In: *Applied Energy* 267 (June 2020), p. 115100. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115100.

[68] Yacine Alimou et al. "Assessing the security of electricity supply through multi-scale modeling: The TIMES-ANTARES linking approach". en. In: *Applied Energy* 279 (Dec. 2020), p. 115717. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115717.

[69] M. Geidl and G. Andersson. "A modeling and optimization approach for multiple energy carrier power flow". In: *2005 IEEE Russia Power Tech*. June 2005, pp. 1–7. DOI: 10.1109/PTC.2005.4524640.

[70] Ivalin Petkov and Paolo Gabrielli. "Power-to-hydrogen as seasonal energy storage: an uncertainty analysis for optimal design of low-carbon multi-energy systems". en. In: *Applied Energy* 274 (Sept. 2020), p. 115197. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115197.

[71] Eurostat. *Electricity prices by type of user*. 2021.

[72] Iain Dunning, Joey Huchette, and Miles Lubin. "JuMP: A Modeling Language for Mathematical Optimization". en. In: *SIAM Review* 59.2 (Jan. 2017), pp. 295–320. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/15M1020575.

[73] Arthur Clerjon and Fabien Perdu. "Matching intermittency and electricity storage characteristics through time scale analysis: an energy return on investment comparison". en. In: *Energy & Environmental Science* 12.2 (2019), pp. 693–705. DOI: 10.1039/C8EE01940A.

[74] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. "Design of distributed energy systems under uncertainty: A two-stage stochastic programming approach". en. In: *Applied Energy* 222 (July 2018), pp. 932–950. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.04.019.

[75] Alan J. King. *Modeling with stochastic programming*. en. Springer series in operations research and financial engineering. New York: Springer, 2012. ISBN: 978-0-387-87816-4.

[76] Tristan Rigaut et al. "Stochastic Optimization of Braking Energy Storage and Ventilation in a Subway Station". In: *IEEE Transactions on Power Systems* 34.2 (Mar. 2019), pp. 1256–1263. ISSN: 1558-0679. DOI: 10.1109/TPWRS.2018.2873919.

[77] Tao Hong et al. "Energy Forecasting: A Review and Outlook". In: *IEEE Open Access Journal of Power and Energy* 7 (2020), pp. 376–388. ISSN: 2687-7910. DOI: 10.1109/OAJPE.2020.3029979.

[78] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. "A review of uncertainty characterisation approaches for the optimal design of distributed energy systems". en. In: *Renewable and Sustainable Energy Reviews* 88 (May 2018), pp. 258–277. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.02.021.

[79] Stefan Pfenninger and Iain Staffell. "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data". en. In: *Energy* 114 (Nov. 2016), pp. 1251–1265. ISSN: 0360-5442. DOI: 10.1016/j.energy.2016.08.060.

[80] Kumar Biswajit Debnath and Monjur Mourshed. "Forecasting methods in energy planning models". en. In: *Renewable and Sustainable Energy Reviews* 88 (May 2018), pp. 297–325. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.02.002.

[81] Sandhya Patidar, David P. Jenkins, and Sophie A. Simpson. "Stochastic modelling techniques for generating synthetic energy demand profiles". In: *International Journal of Energy and Statistics* 04.03 (Sept. 2016), p. 1650014. ISSN: 2335-6804. DOI: 10.1142/S2335680416500149.

[82] Stylianos I. Vagropoulos et al. "ANN-based scenario generation methodology for stochastic variables of electric power systems". en. In: *Electric Power Systems Research* 134 (May 2016), pp. 9–18. ISSN: 0378-7796. DOI: 10.1016/j.epsr.2015.12.020.

[83] Yize Chen et al. "Model-Free Renewable Scenario Generation Using Generative Adversarial Networks". In: *arXiv:1707.09676 [cs, math]* (Feb. 2018). arXiv: 1707.09676.

[84] Weicong Kong et al. "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network". In: *IEEE Transactions on Smart Grid* 10.1 (Jan. 2019), pp. 841–851. ISSN: 1949-3061. DOI: 10.1109/TSG.2017.2753802.

[85] Iman Ghalehkhondabi et al. "An overview of energy demand forecasting methods published in 2005–2015". en. In: *Energy Systems* 8.2 (May 2017), pp. 411–447. ISSN: 1868-3975. DOI: 10.1007/s12667-016-0203-y.

[86] Drury B. Crawley et al. "EnergyPlus: Energy Simulation Program". In: *ASHRAE Journal* 42 (2000), pp. 49–56.

[87] University of Wisconsin–Madison. Solar Energy Laboratory. *TRNSYS, a transient simulation program*. Madison, Wis. : The Laboratory, 1975., 1975.

[88] Thomas Berthou et al. "Development and validation of a gray box model to predict thermal behavior of occupied office buildings". en. In: *Energy and Buildings* 74 (May 2014), pp. 91–100. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2014.01.038.

[89] Francesco Lombardi et al. "Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model". en. In: *Energy* 177 (June 2019), pp. 433–444. ISSN: 0360-5442. DOI: 10.1016/j.energy.2019.04.097.

[90] Jimeno A. Fonseca and Arno Schlueter. "Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts". en. In: *Applied Energy* 142 (Mar. 2015), pp. 247–265. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2014.12.068.

[91] "A verification of CitySim results using the BESTEST and monitored consumption values". In: *Proceedings of the 2nd Building Simulation Applications conference* (2015). Ed. by Emmanuel Walter and Jérôme Henri Kämpf.

[92] Erich Schubert and Peter J. Rousseeuw. "Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms". en. In: *Similarity Search and Applications*. Ed. by Giuseppe Amato et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 171–187. ISBN: 978-3-030-32047-8. DOI: 10.1007/978-3-030-32047-8_16.

[93] A. Saltelli, ed. *Global sensitivity analysis: the primer*. en. OCLC: ocn180852094. Chichester, England ; Hoboken, NJ: John Wiley, 2008. ISBN: 978-0-470-05997-5.

[94] Soha Saad. "Méthodologie de réorganisation du trafic ferroviaire par analyse de sensibilité régionale : application à un incident sur infrastructure électrique". fr. PhD thesis. Université Grenoble Alpes, Oct. 2019.

[95] Stein W. Wallace. "Decision Making Under Uncertainty: Is Sensitivity Analysis of Any Use?" en. In: *Operations Research* 48.1 (Feb. 2000), pp. 20–25. ISSN: 0030-364X, 1526-5463. DOI: 10.1287/opre.48.1.20.12441.

[96] John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. en. 2nd ed. Springer Series in Operations Research and Financial Engineering. New York: Springer-Verlag, 2011. ISBN: 978-1-4614-0236-7. DOI: 10.1007/978-1-4614-0237-4.

[97] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust Optimization*. Princeton University Press, 2009. ISBN: 978-0-691-14368-2.

[98] Samaneh Pazouki and Mahmoud-Reza Haghifam. “Optimal planning and scheduling of energy hub in presence of wind, storage and demand response under uncertainty”. en. In: *International Journal of Electrical Power & Energy Systems* 80 (Sept. 2016), pp. 219–239. ISSN: 0142-0615. DOI: 10.1016/j.ijepes.2016.01.044.

[99] Giovanni Micheli et al. “A two-stage stochastic MILP model for generation and transmission expansion planning with high shares of renewables”. en. In: *Energy Systems* (Oct. 2020). ISSN: 1868-3975. DOI: 10.1007/s12667-020-00404-w.

[100] Francisco D. Munoz and Jean-Paul Watson. “A scalable solution framework for stochastic transmission and generation planning problems”. en. In: *Computational Management Science* 12.4 (Oct. 2015), pp. 491–518. ISSN: 1619-6988. DOI: 10.1007/s10287-015-0229-y.

[101] Stefano Moret et al. “Decision support for strategic energy planning: a robust optimization framework”. In: *European Journal of Operational Research* (June 2019). ISSN: 0377-2217. DOI: 10.1016/j.ejor.2019.06.015.

[102] Stefano Moret et al. “Overcapacity in European power systems: Analysis and robust optimization approach”. en. In: *Applied Energy* (Dec. 2019), p. 113970. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.113970.

[103] Paolo Gabrielli et al. “Robust and optimal design of multi-energy systems with seasonal storage through uncertainty analysis”. en. In: *Applied Energy* 238 (Mar. 2019), pp. 1192–1210. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.01.064.

[104] Xu Andy Sun and Antonio J. Conejo. *Robust Optimization in Electric Energy Systems*. en. Vol. 313. International Series in Operations Research & Management Science. Cham: Springer International Publishing, 2021. ISBN: 978-3-030-85127-9 978-3-030-85128-6. DOI: 10.1007/978-3-030-85128-6.

[105] Zhe Zhou et al. “A two-stage stochastic programming model for the optimal design of distributed energy systems”. en. In: *Applied Energy* 103 (Mar. 2013), pp. 135–144. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2012.09.019.

[106] Masoud Sharafi and Tarek Y. ElMekkawy. “Stochastic optimization of hybrid renewable energy systems using sampling average method”. en. In: *Renewable and Sustainable Energy Reviews* 52 (Dec. 2015), pp. 1668–1679. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.08.010.

[107] J. Dupačová, N. Gröwe-Kuska, and W. Römisch. “Scenario reduction in stochastic programming”. en. In: *Mathematical Programming* 95.3 (Mar. 2003), pp. 493–511. ISSN: 1436-4646. DOI: 10.1007/s10107-002-0331-0.

[108] Holger Heitsch and Werner Römisch. “Scenario Reduction Algorithms in Stochastic Programming”. en. In: *Computational Optimization and Applications* 24.2 (Feb. 2003), pp. 187–206. ISSN: 1573-2894. DOI: 10.1023/A:1021805924152.

[109] Davide Fioriti et al. “Comparison among deterministic methods to design rural mini-grids: effect of operating strategies”. In: *2019 IEEE Milan PowerTech*. June 2019, pp. 1–6. DOI: 10.1109/PTC.2019.8810717.

[110] G. Darivianakis et al. “A Data-Driven Stochastic Optimization Approach to the Seasonal Storage Energy Management”. In: *IEEE Control Systems Letters* 1.2 (Oct. 2017), pp. 394–399. ISSN: 2475-1456. DOI: 10.1109/LCSYS.2017.2714426.

[111] Gijs J. H. de Goeijen, Gerard J. M. Smit, and Johann L. Hurink. “Improving an Integer Linear Programming Model of an Ecovat Buffer by Adding Long-Term Planning”. en. In: *Energies* 10.12 (Dec. 2017), p. 2039. DOI: 10.3390/en10122039.

[112] R. Tyrrell Rockafellar and Stanislav Uryasev. “Conditional value-at-risk for general loss distributions”. en. In: *Journal of Banking & Finance* 26.7 (July 2002), pp. 1443–1471. ISSN: 0378-4266. DOI: 10.1016/S0378-4266(02)00271-6.

[113] Rémy Rigo-Mariani et al. “Comparison of optimization frameworks for the design of a multi-energy microgrid”. en. In: *Applied Energy* 257.113982 (Jan. 2020). ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.113982.

[114] Teemu Räsänen and Mikko Kolehmainen. “Feature-Based Clustering for Electricity Use Time Series Data”. en. In: *Adaptive and Natural Computing Algorithms*. Ed. by Mikko Kolehmainen, Pekka Toivanen, and Bartlomiej Beliczynski. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2009, pp. 401–412. ISBN: 978-3-642-04921-7. DOI: 10.1007/978-3-642-04921-7_41.

[115] T. Warren Liao. “Clustering of time series data—a survey”. en. In: *Pattern Recognition* 38.11 (Nov. 2005), pp. 1857–1874. ISSN: 00313203. DOI: 10.1016/j.patcog.2005.01.025.

[116] Adrien Le Franc et al. “EMSx: a numerical benchmark for energy management systems”. en. In: *Energy Systems* (Feb. 2021). ISSN: 1868-3975. DOI: 10.1007/s12667-020-00417-5.

[117] Georgios Mavromatidis and Ivalin Petkov. “MANGO: A novel optimization model for the long-term, multi-stage planning of decentralized multi-energy systems”. en. In: *Applied Energy* 288 (Apr. 2021), p. 116585. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2021.116585.

[118] Zachary K. Pecenak, Michael Stadler, and Kelsey Fahy. “Efficient multi-year economic energy planning in microgrids”. en. In: *Applied Energy* 255 (Dec. 2019), p. 113771. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.113771.

[119] Gauthier Limpens et al. “EnergyScope TD: A novel open-source model for regional energy systems”. In: *Applied Energy* 255 (Dec. 2019), p. 113729. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.113729.

[120] Jingdong Wei et al. "Multi-period planning of multi-energy microgrid with multi-type uncertainties using chance constrained information gap decision method". en. In: *Applied Energy* 260 (Feb. 2020), p. 114188. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.114188.

[121] Angela Flores-Quiroz and Kai Strunz. "A distributed computing framework for multi-stage stochastic planning of renewable power systems with energy storage as flexibility option". en. In: *Applied Energy* 291 (June 2021), p. 116736. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2021.116736.

[122] Markus Bohlayer et al. "Multi-period investment pathways - Modeling approaches to design distributed energy systems under uncertainty". en. In: *Applied Energy* 285 (Mar. 2021), p. 116368. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.116368.

[123] Pierre Haessig, Hamid Ben Ahmed, and Bernard Multon. "Energy storage control with aging limitation". en. In: *2015 IEEE Eindhoven PowerTech*. Eindhoven, Netherlands: IEEE, June 2015, pp. 1–6. ISBN: 978-1-4799-7693-5. DOI: 10.1109/PTC.2015.7232683.

[124] Roman Le Goff Latimier. "Gestion et dimensionnement d'une flotte de véhicules électriques associée à une centrale photovoltaïque: co-optimisation stochastique et distribuée". en. PhD thesis. Université Paris-Saclay, ENS Cachan, 2016.

[125] Davide Fioriti et al. "Multi-year stochastic planning of off-grid microgrids subject to significant load growth uncertainty: overcoming single-year methodologies". en. In: *Electric Power Systems Research* 194 (May 2021), p. 107053. ISSN: 0378-7796. DOI: 10.1016/j.epsr.2021.107053.

[126] Tom Lambert, Paul Gilman, and Peter Lilienthal. "Micropower System Modeling with Homer". en. In: *Integration of Alternative Sources of Energy*. Ed. by Felix A. Farret and M. Godoy Simões. Hoboken, NJ, USA: John Wiley & Sons, Inc., Apr. 2006, pp. 379–418. ISBN: 978-0-471-75562-3 978-0-471-71232-9. DOI: 10.1002/0471755621.ch15.

[127] EDF. *Electricité - Tarif Bleu EDF : Option Base ou Heures Creuses*. 2020.

[128] N Lebedeva et al. *Li-ion batteries for mobility and stationary storage applications: scenarios for costs and market growth*. en. Tech. rep. OCLC: 1111111915. European commission, Joint Research Center, 2018.

[129] IRENA. *Renewable power generation costs in 2018*. en. Tech. rep. Abu Dhabi: International Renewable Energy Agency, 2019, p. 88.

[130] A. Hina Fathima and K. Palanisamy. "Optimization in microgrids with hybrid energy systems – A review". en. In: *Renewable and Sustainable Energy Reviews* 45 (May 2015), pp. 431–446. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.01.059.

[131] Emilio L. Cano et al. "A strategic optimization model for energy systems planning". en. In: *Energy and Buildings* 81 (Oct. 2014), pp. 416–423. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2014.06.030.

[132] Atom. *Atom Flight Manual*.

[133] Openmod. *Open Energy Modelling Initiative Manifesto*. 2018.

[134] Henrik Lund et al. "EnergyPLAN – Advanced analysis of smart energy systems". en. In: *Smart Energy* 1 (Feb. 2021), p. 100007. ISSN: 2666-9552. DOI: 10.1016/j.segy.2021.100007.

[135] Mark Howells et al. "OSeMOSYS: The Open Source Energy Modeling System: An introduction to its ethos, structure and development". en. In: *Energy Policy*. Sustainability of biofuels 39.10 (Oct. 2011), pp. 5850–5870. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2011.06.033.

[136] Tom Brown, Jonas Hörsch, and David Schlachtberger. "PyPSA: Python for Power System Analysis". In: *Journal of Open Research Software* 6 (Jan. 2018). arXiv: 1707.09913, p. 4. ISSN: 2049-9647. DOI: 10.5334/jors.188.

[137] Stefan Pfenninger and Bryn Pickering. "Calliope: a multi-scale energy systems modelling framework". en. In: *Journal of Open Source Software* 3.29 (Sept. 2018), p. 825. ISSN: 2475-9066. DOI: 10.21105/joss.00825.

[138] Jeff Bezanson et al. "Julia: A Fresh Approach to Numerical Computing". In: *SIAM Review* 59.1 (Jan. 2017), pp. 65–98. ISSN: 0036-1445. DOI: 10.1137/141000671.

[139] Modelica. *Modelica® - A unified Object-Oriented Language for Systems Modeling*. 2021.

[140] R. Tyrrell Rockafellar and Stanislav Uryasev. "Optimization of conditional value-at-risk". en. In: *The Journal of Risk* 2.3 (2000), pp. 21–41. ISSN: 14651211. DOI: 10.21314/JOR.2000.038.

[141] Michal Kaut and Stein W Wallace. "Evaluation of Scenario-Generation Methods for Stochastic Programming". en. In: (2003), p. 16.

[142] J. M. Morales et al. "Scenario Reduction for Futures Market Trading in Electricity Markets". In: *IEEE Transactions on Power Systems* 24.2 (May 2009), pp. 878–888. ISSN: 1558-0679. DOI: 10.1109/TPWRS.2009.2016072.

[143] Karl Pearson. "LIII. On lines and planes of closest fit to systems of points in space". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (Nov. 1901), pp. 559–572. ISSN: 1941-5982. DOI: 10.1080/14786440109462720.

[144] Leland McInnes, John Healy, and James Melville. “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction”. In: *arXiv:1802.03426 [cs, stat]* (Sept. 2020). arXiv: 1802.03426.

[145] Martin Ester et al. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”. In: *KDD*. 1996.

[146] “Dynamic Time Warping”. en. In: *Information Retrieval for Music and Motion*. Ed. by Meinard Müller. Berlin, Heidelberg: Springer, 2007, pp. 69–84. ISBN: 978-3-540-74048-3. DOI: 10.1007/978-3-540-74048-3_4.