Sizing and management of hybrid renewable energy system for data center supply

Maroua Haddad

UBFC UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

UFC UNIVERSITÉ DE FRANCHE-COMTÉ

THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ PRÉPARÉE À L'UNIVERSITÉ DE FRANCHE-COMTÉ

École doctorale n°37

Sciences Pour l'Ingénieur et Microtechniques

Doctorat d'Informatique

par

MAROUA HADDAD

Sizing and management of a hybrid renewable energy system for data center supply

Dimensionnement et gestion d'un système hybride d'énergie renouvelable pour l'alimentation d'un data center

Thèse présentée et soutenue à Besançon, le 28 November 2019

Composition du Jury :

LECUIVRE JÉRÔME	Docteur Ingénieur de recherche chez EATON à Grenoble	Examinateur
LEFÈVRE LAURENT	Chargé de Recherche à l'INRIA à l'ENS de Lyon	Rapporteur
NICOD JEAN-MARC	Professeur à l'ENSMM à Besançon	Directeur de thèse
PÉRA MARIE CÉCILE	Professeur à l'Université de Franche-Comté à Belfort	Co-directrice de thèse
PIERSON JEAN-MARC	Professeur à l'Université Paul Sabatier à Toulouse	Examinateur
ROCHE ROBIN	Maître de Conférences à l'UTBM à Belfort	Examinateur
VARNIER CHRISTOPHE	Maître de Conférences à l'ENSMM à Besançon	Invité
VINASSA JEAN-MICHEL	Professeur à Bordeaux INP	Examinateur
WURTZ FRÉDÉRIC	Directeur de Recherche au CNRS à Grenoble INP	Rapporteur

N° X X X

SPIM

école doctorale **sciences pour l'ingénieur et microtechniques**

Title: Sizing and management of a hybrid renewable energy system for data center supply

Keywords: Hybrid energy system, off-grid, Optimization, Renewable energy sources, photovoltaic panels, wind turbines, standalone, Energy management, Sizing, Energy storage, linear programming, green data centers, Scheduling.

Abstract:

Information and communication technologies have recently become a major sector in energy consumption, particularly with the advent of large platforms on the Internet. These platforms use data centers, which concentrate a very large number of machines processing information and providing services, causing a high energy consumption. The use of renewable energy sources (RES) on-site is then a promising way to reduce their ecological impact. However, some renewable energies such as solar and wind energy are intermittent and uncertain, being related to weather conditions. Since a data center must maintain a certain quality of service, using these sources effectively requires the usage of storage devices. This thesis explores an efficient sizing and management methods for a hybrid renewable energy infrastructure composed of wind turbines, photovoltaic panels, batteries and a hydrogen system..

A first contribution addresses the problem of sizing the electrical plateform in order to meet the data center demand. A sizing tool is proposed, taking several metrics into account and providing three different system configurations as solutions. The user therefore chooses an appropriate configuration, according to his global economic plan of his H_2 ecosystem. A second contribution studies the problem of energy management using a mixed integer linear programming approach. An optimal management tool is therefore provided to find various source schedules according to different user's objectives. The obtained solutions are discussed with several metrics considering different time horizon in order to find the best storage management to meet the data center requests. Finally, a third contribution aims to forecast the weather data to obtain a preciser sizing of the sources using SARIMA model in order to reduce forecasts errors.

Titre : Sizing and management of a hybrid renewable energy system for data center supply

Mots-clés : Système d'énergie hybride, Optimisation, Sources d'énergie renouvelables, Gestion de l'énergie, Dimensionnement, Stockage d'énergie, Programmation linéaire, Centres de données verts

Résumé :

Le secteur du numérique est récemment devenu un secteur majeur de la consommation d'electricité dans le monde, notamment avec l'avènement des data centers qui concentrent un trés grand nombre de machines traitant des informations et fournissant des services. L'utilisation de sources d'énergie renouvelables sur site est un moyen prometteur de réduire l'impact écologique des data centers. Cependant, certaines énergies renouvelables comme les énergies solaire et éolienne sont intermittentes, étant liées aux conditions météorologiques. Étant donné qu'un centre de données doit maintenir une certaine qualité de service, l'utilisation efficace de ces sources nécessite l'utilisation de stockages. Cette thèse explore à la fois une méthode dimensionnement et une méthode de gestion optimale d'une infrastructure hybride d'énergie renouvelable, composée de panneaux photovoltaïques, d'éoliennes, de batteries et de système de stockage hydrogène.

Une première contribution aborde le problème du dimensionnement de cette infrastructure électrique afin de répondre à la demande du data center. Un outil de dimensionnement est proposé, prenant en compte plusieurs métriques et fournissant trois configurations différentes. L'utilisateur choisit donc la configuration approprié, en fonction de son plan économique global de son écosystème H_2. Une deuxième contribution étudie le problème de la gestion de l'énergie par programmation linéaire en nombres entiers. Un outil de gestion optimal est fourni pour trouver différents engagements optimaux des sources en fonction des objectifs de l'utilisateur. Les solutions obtenues sont ensuite discutées avec plusieurs métriques et avec différents horizons temporelles afin de trouver la meilleure solution pour répondre à la demande du data center. Enfin, une troisième contribution vise à prévoir l'evolution temporelle de l'ensoleillement et de la vitesse du vent à gros grain pour obtenir un dimensionnement plus précis à l'aide du modèle SARIMA.

UBFC UNIVERSITE BOURGOGNE FRANCHE-COMTÉ

Université Bourgogne Franche-Comté
32, avenue de l'Observatoire
25000 BesanÃğon, France

Success is not final, failure is not fatal: it is the courage to continue that counts.

- Winston Churchill

Acknowledgements

I would like to express my sincere gratitude to all the people who contributed to the work presented in this thesis taking part of the ANR DATAZERO project.

First of all, I want to express my deep and sincere gratitude to my engineering diploma supervisor, Pr. Abdelghani Adnane, for supporting me and giving me the opportunity to work at FEMTO-st institute. Without his encouragement and wise words, I would have might missed such an adventure and greet meetings.

Second, I want to express my sincere thanks to my thesis director, Pr. Jean-Marc Nicod for providing valuable guidance and encouragement throughout this research. Thank you for all the things you brought to this work and for your always interesting ideas and comments. Thank you for your ability of orienting me towards the correct direction in my research, for your discernment and vision to the analysis of my results.

Moreover, I am very grateful to my co-supervisor, Pr. Marie-Cécile Péra, for the valuable discussion, her guidance, and her support to work. Thank you for your energy and the time devoted to this work, always ready to help me.

During these three amazing years, I learned how to overcome research difficulties thanks to both of you. I really enjoyed working with you, my supervisor and co-supervisor, thank you once again, I would always be grateful.

I also would like to thank the reviewers for accepting to serve as the reviewers of my thesis and all the members of the committee for the energy devoted to the reading and correction of my document. I am very grateful for their valuable discussions, suggestions, and comments.

My time in Besançon was made memorable and enjoyable in large part due to the friends and groups that became part of my life. I am grateful for the time spent with labmates and friends, for memorable trips into the mountains, for parties and for many other memories.

I cannot finish without revealing my thankfulness and gratefulness to all my family. To my parents you are my best back up team in time of need. Your encouragement makes me overcome all difficulties and get stronger. I am really thankful for having such a beloved parents believing in me even when I don't.

I want to express my deepest gratitude to my beloved brother and sister whose endless love, unconditional support, and guidance are with me in whatever I pursue. I also thank and appreciate my lovely little sister for her unconditional support especially her "night calls".

To Ibrahim, thank you for being patient, supporting me and encouraging me in all moments, even in the worst ones. Thank you for loving me when I was not able to love myself. I will never forget your presence at my side and your support during these years.

Thank you all from the bottom of my heart for everything.

Finally, I would like to sincerely thank everyone that I met here in France. It is an amazing, unforgettable, and valuable experience in my life.

Contents

List of Figures

List of Tables

Introduction

The enormous growth of Internet, the increase in the number of online services accessible to public have highlighted the cost imposed by the Cloud in general and by data centers in particular in terms of energy consumption over the past few years. In fact, due to their high computing performance, data center are getting called factories of the digital age. Actually, a typical data center includes hundreds of thousands of servers, cooling equipment, and substation transformers which contribute hardly in global warming as they deliver a huge quantity of CO2.

According to the Environment and Energy Control Agency, sending an e-mail with a file as an attachment requires the cost of operating a high-powered bulb for one hour. Every day, 10 billion emails are sent around the world, which corresponds on average to 50 GWh, means the electrical production of 15 nuclear power stations operating for one hour. According to the estimates of the year 2013, data-centers in the United States consumed around 91 billions of kWh. European data center consumed 56 billions of kWh in the same year. Consequently, the global electricity part dedicated to their consumption has reached unprecedented levels whee they reached 4 % of the global energy consumption. This electricity consumption influences on climate change as most of the electricity is produced from fossil fuels. A 2008 study has estimated that data-centers around the world emitted 116 million tons of carbon, a little more than what Nigeria emitted in the same year. The cost of electricity is therefore a significant burden on data-center operators, which has led to many green energy initiatives.

In this context, several initiatives are being taken to make the data centers IT components more virtuous in terms of their ecological footprint. Many researchers addressed this problem under different approaches where they focused on using dynamic voltage and frequency scaling (DVFS) for instance, or on consolidation, i.e., clustering virtual machines to avoid powering on too many physical hosts. The reduction of energy consumption however tends to have a limited impact, since it decreases performance. It may even lead to rebound effects as users tend to use more computing resources if they are cheaper. It is thus not an adequate solution to satisfy the ever growing need for computing power while reducing the carbon emissions.

Another approach able to limit the world-wild alert relies on energy sources that matches with the present day demand for clean, affordable, and sustainable energy is the use of renewable energy sources. Moreover, it allows the energetic independence of these critical facilities. In a general sense, renewable energy as it is currently conceived, refers to energy sources that have low environmental footprint such as solar panel or wind turbines. Nevertheless, due to their main characteristic, intermittency, their integration in data centers is however causing technical difficulties.

In fact, real-time balancing of supply and demand is a challenging task. Many criteria must be taken into account in a daily supply of data centers such as reliable supply, quality services, demand fluctuations. that must take into account the needs of data centers for highly reliable. As a consequence, supplying a data center using only renewable is currently unfeasible without adding energy storage or classical generators.

Utilization of these intermittent energy sources in a proper hybrid structure combining them with or without grid connection, but with an energy storage/back-up power production unit (battery, fuel cell, diesel generator, etc.) is the most accepted solution to overcome the mentioned drawbacks of these new generation power production systems especially for electrification of rural areas without investment in grid facilities. Due to the intermittent nature of wind and solar energy, stand-alone wind and PV energy systems normally require energy storage devices or some other generation sources to form a hybrid system.

Then, several companies and researchers has been focusing on this approach in order to make their data center the greenest possible. They can either generate their own renewable energy (self-generation), draw it from a nearby plant (co-location) or simply buy renewable energy from an external supplier on markets or via a power purchase commitment. For example, Apple with the Techren Solar II project will produce 200 MW of solar power for their data center in Reno, Nevada. In parallel, many researchers are working on the use of renewable energy sources and storage units to supply datacenters. The “GreenDataNet” European project has proposed an optimization solution, where they combines solar panels and batteries, that allows urban datacenters to radically improve their energy performance. Similarly, the “RenewIT” project has developed a simulation tool for datacenter operators to select the optimum combination of efficiency measures, and renewables, for energy and carbon sustainable facilities in several European climate regions.

Although, none of these project handle a green data center only supplied by renewable energy. For this reason, the DATAZERO project, funded by the French National research Agency (ANR), adopts an original approach. It consists in computing data center’s resources provisioned with 100% renewable energy. Designing and operating such a data center is however not trivial. There is a need to optimize the IT load to the energy availability, and conversely, to optimize the energy production to the incoming IT load. The sizing of the system in terms of IT and electrical components is also a huge challenge. Thus, the ANR DATAZERO project is divided on two main decision modules; an IT decision module (ITDM) responsible for the management of the IT part (scheduling tasks, virtualization of servers, etc.) and a power decision module (PDM) responsible for the management of the renewable sources (source commitment, storage, etc.). The project ambition consists in converging user needs with renewable energy production thanks to a negotiation loop (NM) [103] as explained in Figure 1.

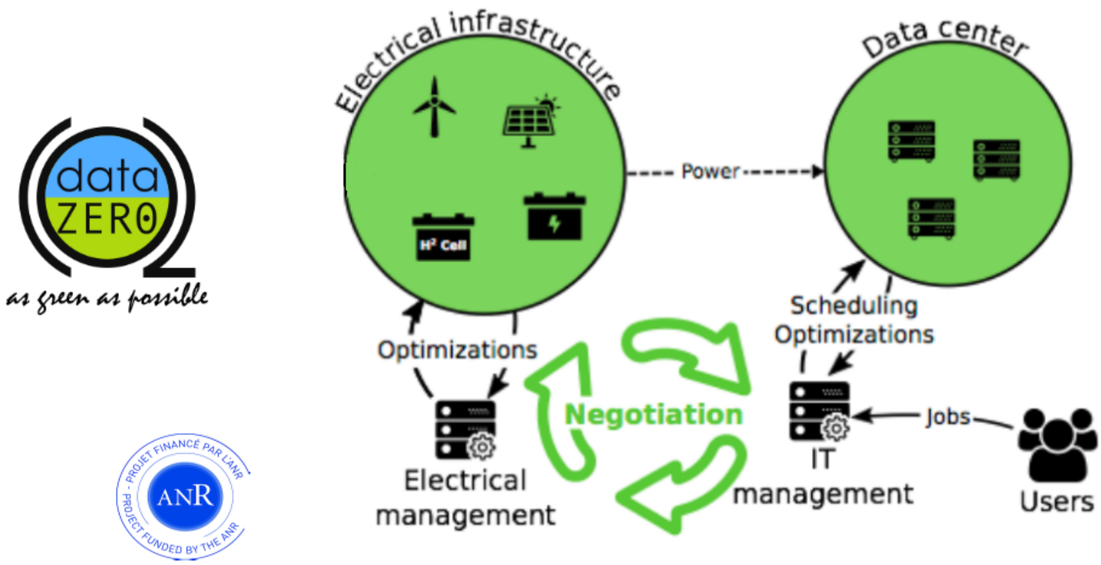


Figure 1: evolution of the renewable energy production by sources

In fact, at each step of this negotiation, one asks to receive one or multiple energy profiles that could be delivered by the data center electrical part and needed to be consumed by the IT part to satisfy a given level of computation for final clients. The negotiation will continue bargaining with the two modules till reaching the achieved convergence between the IT and electrical power profiles.

So, depending on the circumstances and on the period of the year, the level of energy available for the calculations may vary. The principle is that the sizing of both the IT and the electrical parts must allow a computation service in accordance with the quality of service announced in full autonomy thanks to short and long-term storage devices. These storage devices help smooth the energy production available to the data center so that they can provide online services all day and all year long. However, the energy demand is not constant over time because client demand varies. Thus, minimizing the use of storage in view of a given calculation demand to satisfy, and therefore an electrical power to supply, is always the best solution. Therefore, in case of reduction of the energetic envelope, machines can be extinguished, virtual machines can be migrated, etc. Conversely, in the event of a strong demand from the servers, some of the energy can be drawn in addition to that coming directly from the wind turbines and / or solar panels. Here, long-term storage is done through the production of hydrogen with electrolyzers and the reverse generation of electricity is done with fuel cells.

Thus, this thesis focuses on the electrical infrastructure. As part of the ANR DATAZERO project, an efficient sizing and management tools for a hybrid renewable energy infrastructure, composed of photovoltaic panels, wind turbines, hydrogen and battery storage system, is presented. Thus, the chapters in this dissertation are organized as follows:

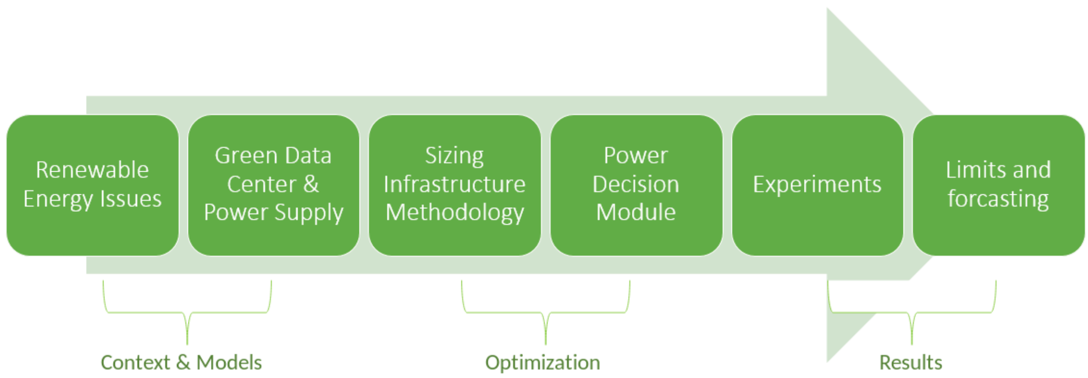


Figure 2: Thesis plan

chapter 1 In this chapter, the issues of the national, European and world wild energy consumption and CO2 emission world wild are presented. As this situation raises a global alert concerning global warming, to solve this problem using renewable energy system presents the most popular solution. Thus, a presentation about these hybrid renewable energy system is given. The chapter ends up by an application context studied in this manuscript from the DATAZERO project and the motivation and objectives.

chapter 2 This chapter presents the techno-economic models, existing and retained for the study of a data center supplied by renewable energy coupled with Energy Storage System (ESS). An overview about data center is presented at first. Then, the models of the chosen system composed of photovoltaic panels, wind turbines, batteries and hydrogen systems. This chapter ends up by a review on different strategies adopted by researchers in designing or managing these systems in order to supply power demand.

chapter 3 This chapter presents the sizing strategy settled in order to design the infrastructure able to supply the data center demand. The chapter starts by detailing the principle of the sizing in order to understand the follow-up ideas. Then the methodology employed is explained. This chapter ends up by treating the results obtained from the sizing tools using common applied performance metric.

chapter 4 This chapter aims to determine different optimal management policy for the hybrid renewable energy system in order to meet several objectives defined the DATAZERO project. The optimization method used is linear programming. Thus, this chapter starts by a definition of the power decision module and its components. Then, the constraint used are explained and linearized in order to construct a mixed integer linear program (MILP). This chapter ends up by playing different scenarios of the mixed integer linear program in order to be able to deliver the power load needed by the negotiation module.

chapter 5 This chapter presents different playing scenarios of the sizing tool and management tool used in the electrical infrastructure presented in the last two chapters. Then, this chapter presents as set a set of scenarios against which the project will be challenged.

chapter 6 As the obtained solutions in the precedent chapters are quite related to the meteorological data downloaded, a valid forecasting of meteorological data need to be studied in order to get the best trends of during the next years. Based on these trends, a sustainable, optimal and green management during the lifetime of the system is ensured. Thus, in this chapter, one forecasting model is defined and used in order to evaluate the robustness of our approach.

Chapter 1

Renewable energy issues

In this chapter, a general introduction is presented for a better understanding of the thesis context. First, the world wild, European and national energy issues are presented. This situation raises a global alert concerning the greenhouse gas (GHG) emissions, air pollution, social concerns and other energy security issues. As a result, many alternative approaches are considered in order to solve these problems. One of the most popular solutions is the use of renewable energy system. Thus, in this chapter, a presentation about these systems and their issues is given. However, due to their intermittent and variable nature, the integration of such sources is a challenge to reach a low-carbon society as it needs storage energy system (SES) to enable satisfaction of the demand. Then, the SES are defined and compared. The chapter ends up by an application context studied in this manuscript from the DATAZERO project and the motivation and objectives

Contents

1.1/ Renewable Intermittent Energy

The renewable intermittent energy projects respond to two major objectives shared at world wild and European level: limiting CO_2 emissions, the main gas responsible for the greenhouse effect and reducing dependence on fossil fuels because of their unavoidable scarcity and increasing cost without speeking of their negative ecologic impact and the speculative geopolitical risk . The challenge is therefore significant since it involves responding to two major threats of Energy of the 21st century: control and security of energy supply (supply/demand balance) and the mitigation of climate change, in particular through lower GHG emissions. These issues are even more important in isolated sites context.

1.1.1/ Energetic Issues

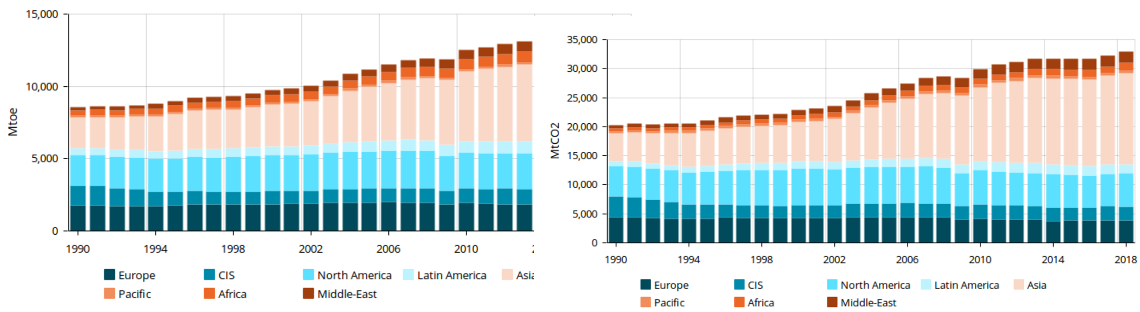
Producing electrical energy, whatever the means used, is a process starting by consuming a primary energy and ending by generating useful electrical energy, with more or less efficiency. This process is classified as follows:

- Primary Energy: are the obtainable energy from nature (e.g., biomass, fossil fuels, wind, etc.), before transformation by human being.
- Secondary Energy: are the conversion results of the primary sources (e.g., gasoline, electricity, etc.). It is also named energy carrier as it allows the user to use it whenever and wherever he wants.
- Final energy: are the delivered energy to the user and directly consumed (e.g., station service gasoline, home electricity). This energy is exchanged through monetary transaction.
- Useful energy: are the transformation of the final energy into services. The efficiency of the user-devices is taken into account (e.g., washing machine, light bulb, etc.).

During this process, some crucial energy losses appears from a transformation to another. These losses prove that energy services can be fulfilled with less primary energy through more efficient conversion processes, delivering processes and devices.

1.1.1.1/ Worldwide energy consumption and emission

For the past decades, the world total primary consumption has been continuously increasing and has almost tripled in 40 years: from 5.000 Million tonnes of oil equivalent (Mtoe) in 1970 to 14.000 Mtoe in 2018 (Figure 1.1a) [33].



(a) Evolution of the energetic wold wild consumption

(b) Evolution of the CO_2 emissions world wild

Figure 1.1: World wild Energetic consumption and emissions. Source: [140]

This corresponds to an average increase of more than 2.2 % per year. In fact, if this fast growth continues, it could reach more than 25000 Mtoe in 2050 [101]. According to the International Energy Agency (IEA), this global consumption should continue to increase in the coming decades

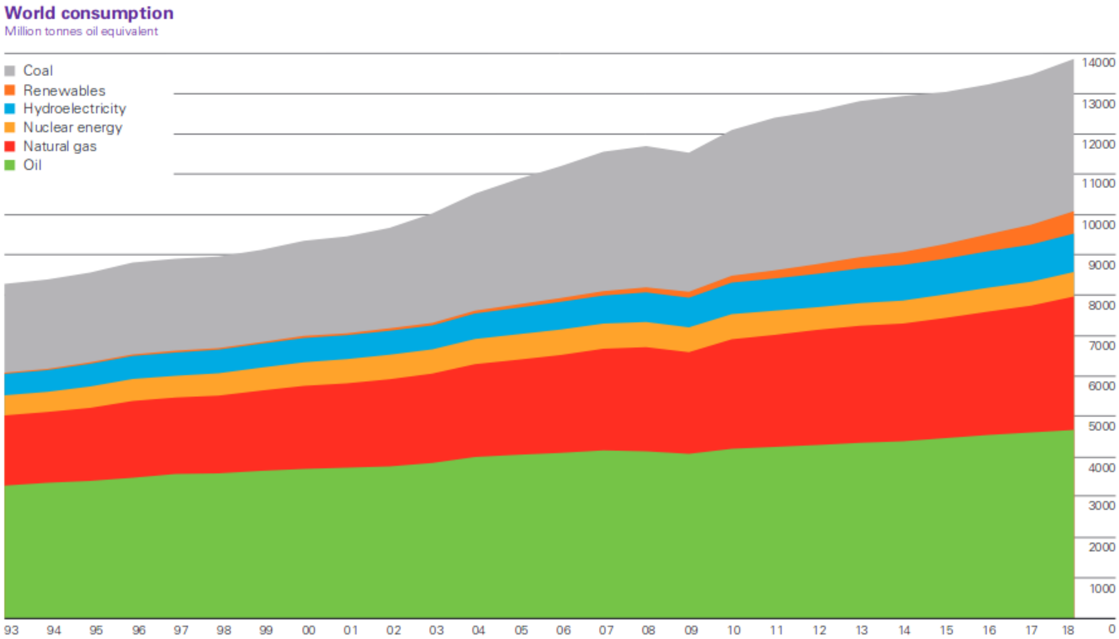


Figure 1.2: Energy sources repartition, consumption and evolution. Source: [33]

because of the growing population (e.g., 7 to 9 billions from 2010 to 2040) and the increasing economic mainly supported by the emerging countries outside OECD[1] like India and China [22].

In 2018, energy combustion emissions count around 80 % of the global carbon dioxide (CO_2) emissions continue to increase (Fig 1.1b). In fact, the IAE has just published its annual report for 2018 [17] pointing out a vigorous growth of CO_2 emissions related to fossil energy, with a 1.7 % increase over 2017. And a historically high volume of 33.1 billion tonnes of CO_2 emitted into the atmosphere.

In 2014, 2015 and 2016, a stagnation followed by two years of slight decrease in emissions has led to believe that the world was following the path advocated by the UN Climate Convention, signed on 1992. Especially that, in December 2015, the Paris Agreement signed by all the countries of the planet confirmed the climatic objectives set on 2009: not to exceed 2° C more than before the industrial revolution on average worldwide.

Figure 1.2 shows that fossil fuels satisfy the vast majority (82 % in 2015) of global needs and, among them, the first source is oil with 34.2 % of supply, then coal (28 %) and gas (22 %). Renewable energies remain stable from 1993 till 2007 and start increasing to reach 10 % of primary energy consumption, of which 8 % for hydropower while nuclear power supplies 5 % of global demand.

The "Grenelle 1"[2] laws in 2009 and "Grenelle 2"[3] in 2010 as well as the law for the energy transition in 2015[4] have reaffirmed, particularly in the perspective of the world climate conference[5] many quantified objectives in terms of efficiency and energy sobriety. Among these objectives, there is the 50 % reduction in energy world wild consumption in 2050 as well as increasing the part of renewable energy (more carbon-free, less polluting energies) to 32 % of final consumption and 40 % of electricity production in 2030.

Indeed, the International Expert Group on Climate (IPCC) recalls in [36] that in addition to the

[1] Organization for Economic Cooperation and Development members (Europe, Australia, Canada, Korea, United states, Israel, Japan, Mexico, New-Zealand)

[2] Law N°2009-967, August 3rd, 2009, related to the implementation of "Le Grenelle de l'environnement".

[3] Law N°2010-788 ,July 12, 2010, on the national commitment to the environment

[4] Law N°2015-992 of 17 August 2015 on the Energy Transition for Green Growth

[5] COP21, Paris, December 2015

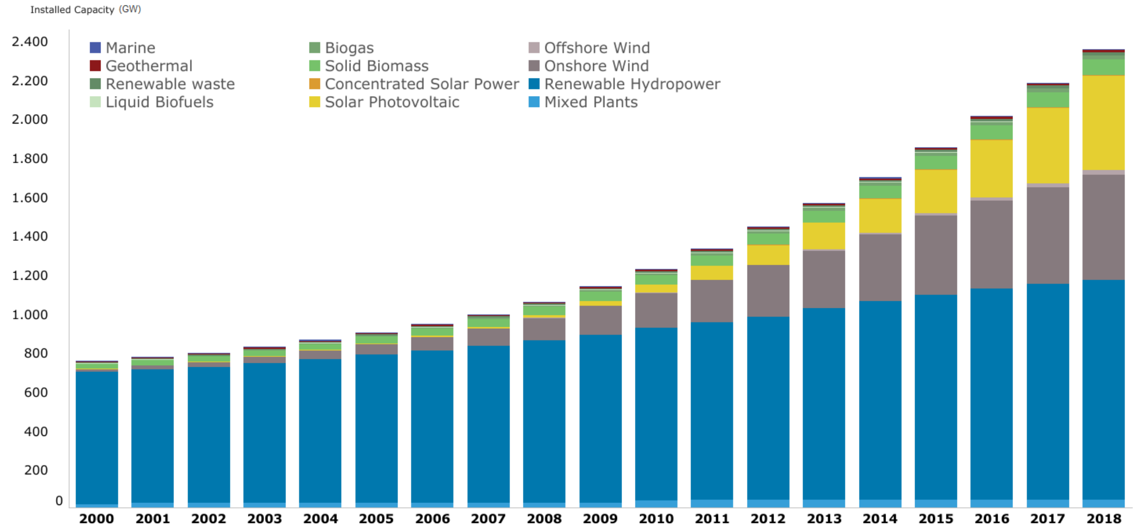


Figure 1.3: evolution of the renewable energy production by sources

initiative of sobriety and energy efficiency, carbon capture and storage processes, renewable energy is one of the essential means of reducing GHG emissions. As a result, this climate challenge is leading to a significant change in the global energy mix. However, this mix must be considered taking into account other major energy issues such as security and control of supply.

1.1.1.2/ Renewable energy

Alternative energy sources are renewable and are thought to be “free” primary energy sources. There are many forms of renewable energy sources including mainly the sun, wind, sea (ocean currents, tides, waves, temperature gradient) and rivers, biomass, as well as the heat of the earth. Then, we speak of solar, wind, hydraulic, wave, tidal, thermal solar, biomass, geothermal. On the other hand, the flows of organic waste and heat from economic activity that can lead to energy recovery are also considered as renewable energies. It is essentially the recovery of waste from agriculture and forestry, the fermentable part of household and industrial waste, energy from industrial processes. RES are produced from these natural sources that are directly inexhaustible on a human scale (wind, sea, sun) but can also be grown as fuelwood or agrofuels. These renewable sources can be converted into energy in various forms to produce useful energy (electricity, heat, cold, fuels):

- Either Grid Tied Systems using power plants (or farms) connected to the high-voltage electricity network or producing heat/cooling distributed by networks;
- Off-Grid Systems (Stand-Alone System): using small-scale installations connected to the low- or medium-voltage electricity network or heat and cold production installations on the scale of one or a few buildings, whether or not connected (site isolated) to the low voltage network.

There are six types of renewable energy [7]:

1. **Wind power:**
   produced from atmospheric winds, either on land (onshore) or on sea (offshore). It exactly comes from atmospheric changes; changes in temperature and pressure makes the air move around the surface of the earth. Wind power is a clean energy source that consist in converting wind energy into electricity by using wind turbines. A wind turbine creates reliable, cost-effective, pollution free energy. It is affordable, clean and sustainable. One wind turbine

can be sufficient to generate enough electrical energy for a household, assuming the location is suitable. Because it is a renewable resource which is non-polluting and renewable, wind turbines create power without using fossil fuels, without producing greenhouse gases or radioactive or toxic waste.

2. **Solar energy:**
produced from solar rays, either from global radiation on photovoltaic cells (PV) to produce electricity or from direct radiation to heat a heat transfer fluid to produce electricity (solar thermodynamic plant, Concentrating Solar Power Plant (CSP)). Installing solar power systems in a residential setting generally means setting up a solar photovoltaic or a solar thermal system on the roof. Photovoltaic means: Photo="light" and photons=energy particles coming from sunlight; voltaic=producing a voltage or volts. Abbreviation=PV. Solar energy is a renewable free source of energy that is sustainable and totally inexhaustible at the humanity time scale, unlike fossil fuels that are finite. It is also a non-polluting source of energy and it does not emit any greenhouse gases when producing electricity. Solar electricity can supplement entire or partial energy consumption.

3. **Marine energy:**
produced from either sea currents, swell or waves, the temperature gradient between surface and depth of the sea, or of the tide.

4. **Energy from biomass:**
produced from the combustion of organic matter (wood, plants, animals, biogas after methanization) or chemical transformations (agrofuels) [124]. It is the second source of renewable energy in France after hydraulics. It is considered renewable if there is sustainable management of the resource limiting environmental impacts. The combustion of wood energy is a strong CO_2 emitter, therefore, carbon capture and storage systems are needed [124, 147]. Energy produced from urban waste in Energy Valuation Units (EVU) is not considered renewable energy. Biomass fuels provided about 5 % of total primary energy use in the United States in 2017. Of these 5 %, about 47 % were from biofuels (mainly ethanol), 44 % were from wood and wood-derived biomass, and 10 % were from the biomass in municipal waste [147]. Plans for biomass data centers are underway in several states. As reported in Environmental Leader in 2018, HP Labs has developed a system that could power a one megawatt data center using manure from a 10,000 cow farm - with enough electricity left over the run of the farm. In Missouri, commity leaders are pushing for a data center to be powered by grass, wood or hay [104].

5. **Hydroelectric power:**
Micro hydro systems, also called hydroelectric power systems convert the flow of water into electrical energy. A turbine can be fully immersed in water. The flowing water rotates the turbine's blades. The amount of energy created depends on the amount of water flowing on the turbine as well as the size of the turbine. Thus, it is produced from dams on rivers or artificial lakes. Large scale hydro power provides about one-quarter of the world's total electricity supply, virtually all of Norway's electricity and more than 40 % of the electricity used in developing countries. The technically usable world potential of large-scale hydro is estimated to be over 2200 GW, of which only about 25 % is currently exploited [27]. Hydropower is by far the leading source of renewable electricity: 3782.0 TWh over 5016.4 TWh, or 75 % in 2013 of global renewable production [32]. The hydraulic sector, with an installed capacity of 25.4 TW, represents 13.8 % of French electricity consumption. There are two small-scale hydropower systems: micro hydropower systems (MHP) with capacities below 100 kW and small hydropower systems (SHP) with capacities between 101 kW and 1 MW. In the developing countries, considerable potential still exists, but large hydropower projects may face financial, environmental, and social constraints [53].

6. **Geothermal energy:**
exploits the temperature gradient of the earth's crust to produce electricity or heat, in three different ways: the geothermal energy at medium and low level from 30 to 150° C aquifers (usable only for heat), the geothermal energy at very low level exploits the heat

around 10 meters deep for the heating (or air conditioning) of individual dwellings and deep geothermal energy from aquifers or dry rocks from which steam is extracted under high pressure converted into electricity. World Geothermal electricity production remains marginal at 11.7 GWe in 2012, of which nearly a third in the United States [140].

Some renewable energies are considered as stable because they do not vary a lot during their production process such as: geothermal energy, wood energy, hydraulic, tidal, thermal of the seas.

Figure 1.3 represents the world wild evolution of renewable energy productions. It shows a growing share of renewable energy in total production, including hydro, from 18.53 % in 2002 to 20.78 % in 2012 [32].

1.1.1.3/ Issues of renewable energy integration

Intermittent RES are wind, photovoltaic or marine energy with the following characteristics [38]:

1. **Non-programmable:** The level of production can not be adapted to the users consumption profile. The overall output of an electrical system must constantly adjust to consumption. These energies do not allow these adjustments, their production profile is sustained and not scalable.

2. **Unsecured:** these energies are by nature intermittent: we can not unfortunately guarantee the presence of sun or wind. If the production of these energies replaces fossil fuels, it does not spare the investment by ensuring the power supply. Thus, other guaranteed resources of production are therefore necessary to compensate the lack of production (no wind, no sun).

3. **Difficult to predict:** these energies are difficult to predict even if many works are underway in these areas to promote the development of these energies and thus, a better management of the electrical systems. The forecast horizons sought today are essentially of two types:

   - Forecast of the day for the next day to optimize the production planning of the different energy resources used.
   - Intra-day forecasts in order to anticipate potential breaks to supply the demand by adjusting the operating program of the various production facilities.

   There are other less restrictive forecasting needs for the management of the electrical system: forecasts by the week, by the month, to forecast the best fuel supply, effective planning of the maintenance of other means of production, optimization of the strategies of management of hydraulic stocks. The absence of these forecasts leads to a reduction in the optimization of the system. Nevertheless, these medium/long-term forecasts represent a lower stake compared to the two types of forecast mentioned above. In general, additional production resources are needed to compensate the forecasting errors.

4. **Fluctuating:** Wind and photovoltaic installations have rapid fluctuations in their production. This instability requires the implementation of compensation means to adjust the supply-demand balance in real time.

5. **Without inertia** : photovoltaic and DC grid interfaced by power electronics do not offer kinetic energy for PV. It tends to increase the speed of variation of the frequency, making it more difficult to maintain network stability.

When the presence of a RES is massive on an electrical system, these characteristics can then quickly be incompatible with the dynamic management of the network.

1.1.2/ Intermittency and forecasting of renewable energy sources

1.1.2.1/ Intermittency

Intermittency is an irregular criteria for an electricity production. It is linked to the flow variations of the mobilized energy sources. Two elements characterize the degree of intermittency [102]:

1. The high variability of the production due to the big number of parameters that are involved in its characterization ;
2. the ease or difficulty with which the intermittent flow can be predicted at the different time scales (forecasted horizons) concerned.

Intermittent energies combine these two characteristics with divers degrees. It is difficult to compare the sources as they have different characteristics (temporal and spatial), also strongly related to the particularities of the considered site [135].

- wind energy: marked randomness and difficult to predict well in advance;
- wave energy: cycles more regular than the wind that generates them;
- tidal energy: regular and predictable cycles;
- solar energy (PV): a priori regular cycles (seasonal and daytime) to which are superimposed variations directly related to the weather (cloudiness) of which we can predict the average value but not the precise value at each moment. This instantaneous value would require predicting the exact nature and timing of cloud passage.

The degree of intermittency also depends on geographical granularity because of the geographic dispersion of the facilities: : wind sites obeying different wind regimes, PV sites sufficiently distant to not be affected by the same cloud systems.

Photovoltaic and wind energies are intermittent energies that are difficult to predict with different typologies of days and high variability. As shown in the examples in Figure 1.3, the power can be relatively high. However, in wind power, the average load factor is, for example, 23 % on the Caribbean Fund in Guadeloupe and 16 % on the site of La Perriere in Reunion. Over a year, the ratio of the duration equivalent to full power on the installed power can vary from single to triple (for example 600h to 1800h for a PV site of Guadeloupe). In addition, the intermittency of RES productions strongly depends on the geographical area on implementation. So, the amounts of energy are relatively low.

On the other hand, the electronic part (related to the connection of certain decentralized productions) should not be neglected as it also impacts the electrical system. In fact, it can strongly impact the stability of the network. Thus, the reduction of the inertia of the network and therefore the increased sensitivity to any disturbance are directly related to the interfacing of these new sources of production and are independent of the intermittent nature of their production. The insertion of RES is a technical and economic challenge as it is necessary to:

- Anticipate brutal events (production drop, ramps of high amplitude) and thus improve the forecast in the short and medium term
- guarantee a reserve of mobilizable power to compensate for the absence of inertia;
- smooth the high frequency intermittencies and make the installations less sensitive.

In general, the aggregation of the sources RES can attenuate the variability of the cumulative production at a very short term (or high frequency) in a limited way, especially for the long term fluctuations [25]. This expansion also depends on the distribution on the territory and on the type of installation. Actions are being taken by both the renewable energy producers and the grid operator

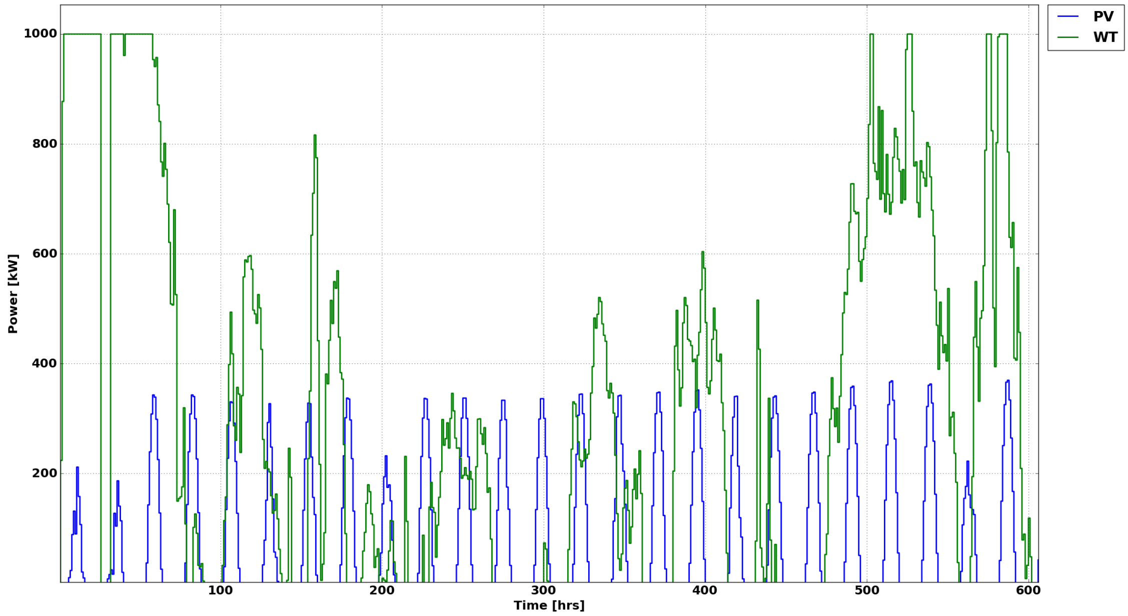


Figure 1.4: Power production variability of the intermittent sources in Los Angeles

to better characterize the output and its intermittency: monitoring of productions in steps of 5 to 10 minutes and measurement campaigns in steps of 1 to 5 seconds, statistical analyzes of the variability, work on forecasting, analysis of impacts on the system.

1.1.2.2/ Forecasting

Detailed information on energy production and consumption forecasts are an essential element for the management and operation strategies of the electricity grid. In fact, reliable forecasts for the production of intermittent renewable energies allow less use of storage and better integration of larger quantities of RES.

Forecasting is an estimate of the short-, medium- or long-term evolution of a variable or phenomenon. It is used in the field of intermittent energies to anticipate future production. Depending on this forecast, decisions can be made and strategies can be applied, particularly concerning the operation of production facilities (RES or storage). The forecast has to be appropriate for the intended application. The forecasting methods are often elaborated in a collaboration between the users (network operators, producers, etc.) and the forecasters in order to define the context and the objectives of their application. As part of intermittent renewable energy participation in the energy mix, the power grid operator needs production forecasts to ensure a high-performance, secure and economical electricity supply system [29].

Two elements characterize the forecast: the forecast horizon and its granularity. The forecast horizon is the time during which the different variables or phenomena are predicted while the granularity corresponds to the time step of this forecast. There are three types of forecast [67]:

- Immediate forecasts: it is a forecast horizon ranging from 15 minutes (m + 15) to 2 hours (H + 2) with a granularity of 30 seconds to 5 minutes. These forecasts make decisions about network regulation and real-time distribution.
- Very short-term forecasts: it is a forecast horizon ranging from 1 hour to 6 hours with hourly granularity. They are used in the load follow and in the planning use update of the sources of production;

- Short-term forecasts: it is a forecast horizon ranging from 1 day to 3 days (D + 1 to D + 3) with hourly granularity. They are used in unit commitment to minimize the overall cost associated to the start-up or shutdown of each sources. These daily forecasts help to optimize the production sources and storage planning, and the control of energy demand. This type of forecast is often made the day before for a production management of the day after D + 1.

In countries with high wind production, forecasting is essential in power grid control systems [92]. In the same way, forecasting the photovoltaic production increasingly occupies an important place, even more in countries where legislation favors the use of solar energy. Many research efforts are currently devoted to improve solar radiation forecasts in order to directly determine the corresponding photovoltaic production forecasts. More reliable solar radiation forecasts will improve the integration of solar RES, either with or without storage, into power grids, particularly in the island.

Determined more and more by specialized companies, forecasting has become a market of interest to both producers and the network operator. However, it remains a technical barrier and a great deal of research work in progress, notably via the use of satellite images [78].

1.2/ Storage energy system (SES)

This section summarizes the issues related to the storage of the intermittent renewable energy (RES) and associated applications. The different storage systems are recalled in Section 1.2.2 but will not be presented in detail in this chapter. The following questions are therefore addressed:

1. Why are we in need to store and for what?
2. How and in what form to store energy, especially from intermittent sources?

1.2.1/ Issues

The challenge of “energy storage” is to provide flexibility and enhance the reliability of energy systems. It is a matter of balancing energy supply and demand, as much for the supply of power as of heat or cold. Today, the answer to consumption peaks; the regulation of power demand during peak hours i.e., high demand. It is mainly achieved either by importing electricity, either by using gas or oil-fired power plants or by means of hydraulic storage. The deployment of other storage systems would not only reduce the cost of imported electricity but also significantly decrease GHG emissions generated by the use of thermal power plants and, in general, by recourse to fossil resources. In addition, the stationary storage of energy for both electricity and thermal storage, is associated to the development of renewable energies as it guaranties a current quality on the distribution network. In fact, the intermittent energy production thanks to solar and/or wind generates significant fluctuations which disturb and deteriorate the distribution equipment. In addition, the intermittent supply is often out of step with the demand, for example, the photovoltaic at the peak demand in the evening. The storage, according to its dimensioning and to its technical specifications, backed by an intermittent source of energy, can offer guarantees in terms of power, predictability, stability, or even provide a programmable energy (“dispatchable”) while providing system services required for the proper functioning of electrical systems [71].

In the current context of renewable energy development, energy storage appears as a possible solution to:

- promote the integration of fluctuating renewable energies,
- improve energy efficiency,
- provide flexibility to energy systems and to enhance the safety of networks.

The ADEME [6], the ATEE [7] and the DGCIS [8] study on the potential of energy storage by 2030. They have confirmed that power storage can play an important role in the supply-demand balance, particularly on cold days [3]. Indeed, by storing at times of lower demand and destocking during the few hours of ultra-peak, investments in production capacity can be avoided. This capacity value of power storage represents about half of the valuations presented and remains independent of the evolution of fuel prices. However, it depends heavily on the discharging duration and the penetration rate of power storage: a "long-term" power storage – several days – allows the demand to be shifted to other days when consumption is lower. Whereas a major deployment of power storage "short term" - a few hours - would not solve a prolonged lack of production capacity, the power storage can not be recharged between two periods of ultra-peak. It is therefore recommended to promote innovative R&D projects to reduce the costs of storage. This should enable to respond in the long term to a major deployment of the intermittent fleet, with satisfactory costs.

The optimal energy mix is energy production but also storage in its various forms, including electricity, heat and hydrogen. In this context, it is also important to consider in particular the energy expenditure needed to manufacture (energy investment) and to recycle the elements of the system. For example, studies on electrochemical batteries, including the results of the INVESTIRE program [114], show that the energy investment of an accumulator is very heavy [113]. In addition, other solutions to better adapt production to consumption exist: a production-driven consumption that could be described as intelligent consumption (incentive pricing, off-peak hours for example, load shedding, etc.). Actions in energy management (MDE) - energy efficiency and sobriety - could also help reduce storage needs and improve carbon balances of energy consumption.

The storage of electrical energy is therefore a major challenge to allow a real penetration of vast naturally dispersed renewable resources, especially those of the sea (wind, waves, currents), intermittent and more or less easily predictable. It is also a solution to increase the safety of grid-connected systems with the possibility of centrally controllable decentralized storage. In this, the DOM-TOM represents a field of experimentation conducive to the development of stationary storage of electricity. Even if the French deposit remains limited (200 to 400 MW), the studied projects (surface CAES, Li-Ion batteries) are profitable for the community and the prospects of deployment on a global level are promising, considering not only the islands but also the regions whose electricity network is not very interconnected [71].

1.2.2/ General description

Storing energy is keeping a quantity of energy for future use. By extension, energy storage also refers to the storage of the material containing the energy. There are two types of applications: stationary (or fixed) storage and embedded (or mobile) storage. We are particularly interested in this part of stationary storage which concerns power production, in particular from RES, connected to the network or isolated site.

Storage is not a power generation facility like any other since the operation of a grid-connected storage system can not be considered alone. Indeed, a single storage is a "production tool" that consumes energy: the technical losses induced by the storage result in an energy balance of the negative system. A storage structure therefore absorbs more energy than it restores. The ratio between energy restored and energy absorbed then defines the overall efficiency of the SSE. In this section we give only the main technical information (operating principles, characteristics, typical applications) of the different large scale HSE technologies. The objective is rather to situate the storage of RES in the general context of SSE. Additional details can be found in [118] or in [1].

[6] Agence de l'Environement et de la maitrise de l'Énergie

[7] Association Technique Énergie Environnement

[8] Direction Générale de la Compétitivité de l'Industrie et des Services

1.2.2.1/ Technologies

Energy Storage Systems (SSE) converts electricity, which is difficult to store directly, into potential, kinetic or chemical energy. There are five physico-chemical categories of stationary storage depending on whether the stored energy is in mechanical, chemical, thermal, electrochemical or electromagnetic form.

1. Storage in a mechanical energy form:

    - Pumped-Hydro Energy storage or PHS:

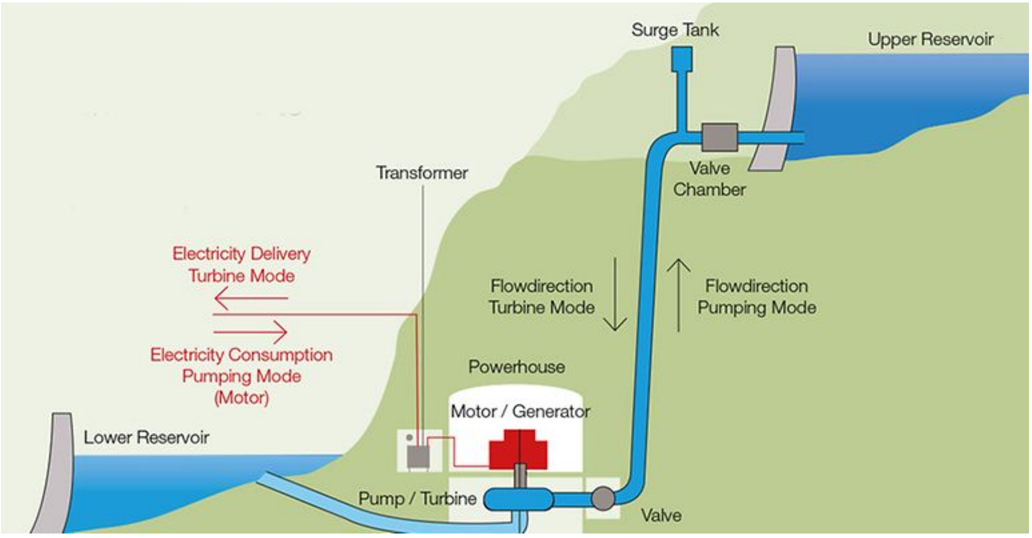


Figure 1.5: Principle of Pumped-Hydro Energy storage. Source [70]

      The PHS can store a large amount of electrical energy through the water potential energy. This type of hydroelectric plant is used to transfer water between two basins at different altitudes. When the network provides surplus electricity, the water from the lower basin is pumped into the upper basin (Figure 1.5).

      Under the effect of gravity, this body of water represents a future power generation capacity. During a deficit of electricity production, the water circulation is reversed: the pump becomes turbine and restores the accumulated energy. With a yield of up to 80 %, it is the most widely used solution for storing energy from power plants [19].

    - Flywheels:
      The energy can be stored as kinetic energy of a wheel-shaped device rotating around its central axis, as shown in Figure 1.6.

      An electric machine provides the kinetic energy (engine operation) and recovers as needed (generator operation), resulting in a decrease in the speed of rotation of the flywheel. The unit operates under vacuum to eliminate frictional energy losses. This system makes it possible to restore more than 80 % of the accumulated energy but for a limited storage time [95]. In practice, the flywheel is used for a very short-term smoothing of energy supply within production devices. This is particularly the case of thermal engines and especially diesel engines.

    - Compressed Air Energy storage or CAES:
      When electricity demand is low, existing systems use old salt mines as reservoirs and a motor-generator-turbine assembly. When electricity demand is high, compressed air in the cavern is used to run a turbine coupled to a generator that produces electricity (Figure 1.7).

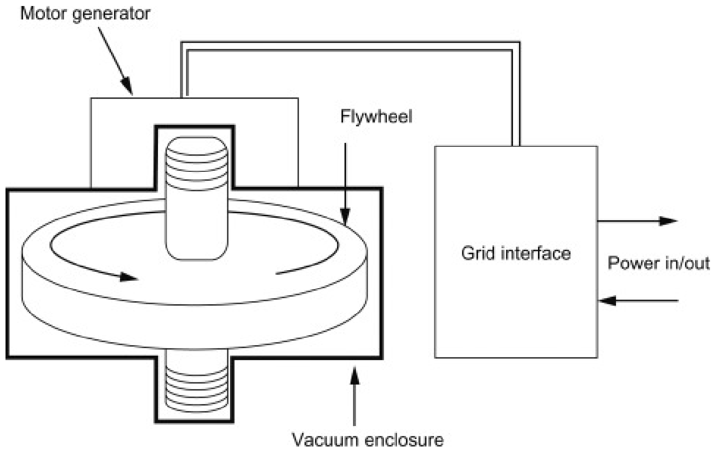


Figure 1.6: Principle of flywheels. Source [14]

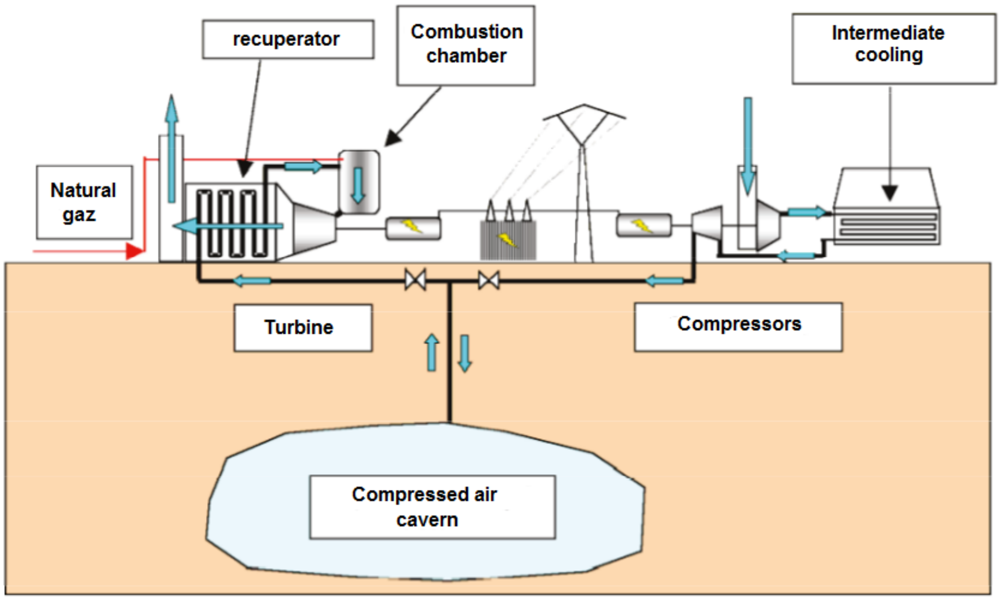


Figure 1.7: Principle of Compressed Air Energy storage [94]

The yield, currently around 50 %, is an axis of research and development. Cave-based compressed air storage from wind and solar power is being piloted in Germany and in the United States.

2. Storage in a chemical energy form:

- Hydrocarbons: Liquid hydrocarbons are currently the dominant form of energy storage in volume, especially for the transport sector. The fuels come from fossil fuels and have a 75 % yield from the "source at the pump", the source – electricity yield being often less than 50 %. Biofuels come from biomass, with a 70 % yield "from biomass to the pump".
- Biomass: Any fuel can be considered as a stock of energy in chemical form. While burning, the compound releases energy in the form of heat that can be recovered and recovered. The term "biomass" refers to all organic materials that can become sources of energy. In the case of plants, it is a form of storage of solar energy: organic matter is derived from CO_2 captured during photosynthesis. However, this energy storage

process is long, of the order of several months, and of low efficiency. Indeed, only 1 % of the solar radiation used during photosynthesis is returned in the form of biomass.

- Hydrogen: does not exist in its natural state but is very abundant on Earth in atomic form H (water, hydrocarbons, etc.). Electrolysis, which consists of breaking down the water molecule into hydrogen and oxygen, requires electricity (cost-effective if electricity production itself is low cost). Hydrogen has the ability to restore energy, making it a particularly attractive form of storage. In fact, hydrogen storage system is composed of an electrolyzer, fuel cell and hydrogen tank where the hydrogen is stored. An **electrolyzer** is a non-reversible device that uses electricity to produce dihydrogen (H2) through the electrolysis process. The process consists in passing a direct electric current through an ionic substance that is either molten or dissolved in a suitable solvent, producing chemical reactions at the electrodes and a decomposition of the materials [74]. According to [30], **fuel cells** are electrochemical devices that convert chemical energy of a fuel, usually hydrogen or hydro-carbon, directly into electricity. Fuel cells consist of an anode, electrolyte, and cathode; and by varying the materials used for each component different fuel cell types are obtained. However, following [37], the most common type is the Proton Exchange Membrane Fuel Cells (PEMFC). The main characteristics of PEMFC are quick start-up, low operating temperatures, and low corrosion. This makes the PEMFC fuel cell a very good option for decentralized power generation.

  Then, electricity produced in excess should be stored for the days with no wind or less sun. The main advantage of hydrogen as energy storage medium is the ability of long term/seasonal storage of different renewable energies in contrast to batteries that can be used for daily storage because of the low storage capacity and the self discharge rate [105].

  One solution is to produce hydrogen through electrolysis [74], splitting with an electric current, of water and to use that hydrogen in a fuel cell to produce electricity during times of low power production or peak demand. When DC current is passed through electrodes immersed in water, hydrogen (at the cathode, negative electrode) and oxygen (at the anode, positive electrode) is generated through electrolysis process. The half reaction at cathode is a reduction ($2H^+(aq) + 2e^- \rightarrow H_2(g)$) and the half reaction at anode is an oxidation ($2H_2O(l) \rightarrow O_2(g) + 4H^+(aq) + 4e^-$). The overall reaction is expressed by:

$$2H_2O \longrightarrow 2H_2 + O_2 \tag{1.1}$$

  The mathematical total energy needed to split the water molecules is defined by the heat of reaction (reverse of the heat of combustion of hydrogen). Conventional values given are either higher heating value (HHV) or lower heating value (LHV), depending on the combustion result if it's liquid water or water vapor. In fact, HHV of a fuel is defined as the amount of heat released by combusting an exact quantity (initially at 25℃) once it is combusted and the products have returned to a temperature of 25℃. In the other sens, LHV is the amount of heat released by combusting a specified quantity (initially at 25℃) and returning the temperature of the combustion products to 150℃.

3. Storage in an electrochemical energy form:

- Electrochemical batteries: The storage of energy in electrochemical batteries is the most common technique for small amounts of electrical energy. Depending on the type of battery (lead-acid, lithium-ion, nickel-metal hybrid, etc.), different chemical reactions are caused by electricity: this is the charging phase of the battery. According to demand, reverse chemical reactions then produce electricity and discharge the system.

  Electrochemical batteries are often intended for portable applications. With a relatively low power, nevertheless, they have a large storage capacity for high discharge times (up to several hours) with a 70 to 90 % efficiency rate. These devices may also have backup functions when the power grid is faulty or in the case of a generation of electricity from renewable energies, with stored energy values of a few Wh up to 40 MWh. The major drawback is related to their lifetime, limited by the chemical degradation of the reactions and their cost.

4. Electrical storage:

- Super Capacitors: Some systems can store energy directly in electrical form. These are mainly super capacitors, electrical components consisting of two conductive armatures that store opposite electrical charges. They are able to deliver a strong power for a very short time. Unlike electrochemical batteries, they can be discharged in a time of about one second or less with a yield of between 80 % and more than 90 %. However, these devices do not store large amounts of energy. Super capacitors have applications in the field of land transport. They ensure in particular the stop and restart at a stop and go.
- Upper conducting magnetic energy storage ou SMES: Another track is that of electromagnetic storage based on superconducting materials. This system is designed to store large amounts of energy, 50 % of which can be restored in less than one second. In addition, such a device has a yield of 75 % to more than 90 %. However, SMES applications, at still very high costs, are still limited and have to demonstrate their feasibility on a large scale, due to the need to maintain a very low temperature. They are mainly developed in the United States.

1.2.2.2/ Comparison

Several solutions are available for energy storage. Nevertheless, they are quite different in their specifications which make them difficult to compare. This is why we have chosen to present a set of technical and economic characteristics that improve the conditions of acceptability of ESS. The Table 1.1 summarizes these characteristics [42]. Flywheels are more suitable for short response times with high discharge powers; they are well established in the markets of critical energy consumption and power control systems (UPS). However, they have significant friction losses that causing a high cost of installation and maintenance [19] . In addition to flywheels, hydrogen systems and circulation batteries are also two prominent technologies in renewable energy storage systems. Hybrid hydrogen/fuel cell photovoltaic production systems are used because they are clean, environmentally friendly, modular and independent of fossil fuels. Circulating batteries allow high power, decoupled power and energy systems, low self-discharge but with low yields compared to other forms of storage technologies. Fuel cells and circulating batteries are suitable for the integration of small-scale renewable energies [95].

	SMES	Super capacitors	Electrochemical batteries	Flywheels	Hydrogen	CAES	PHS
Technical characteristic							
Energy density	1 - 5 Wh/kg	5-10 Wh/kg	20 - 120 Wh/kg	1 - 5 Wh/kg	300 - 600 Wh/kg	8Wh/kg - 200bars	1 kWh/m^3 for a fall of 360m
Energy Form	Magnetic	Electrostatic	Chemical	Mechanical	Chemical	Mechanical	Mechanical
Power			0.1 - 10 MW		1 - 200 kW	till 1000 MW	1 - 100 MW
Time constant	Low: s to 1min	Low: some s to some min	10 min (Nicd) - some 10 h (Pb)	some min to 1h	1h to some days	1h to some days	
Efficiency	> 0.9	> 0.9	0.7 - 0.8	0.9	0.3 - 0.5	0.5	0.65 - 0.8
cyclability	10000 - 100000	10000 - 100000	100 - 5000	10000 - 100000	?	1000 - 10000	> 10000
Reactivity	Ms	ms	ms	ms	?	min	min
Maturity	++	++	+++	+	-/+		+++
Security	Electronic cryogenics of power	No pollutant	heavy metals	blow-out	high re-cyclability		
Economical characteristics							
Cost e/KWh	500 - 72000	50000 - 150000	50 - 200 (Pb) 700 - 2000 (Li)	150 - 25000	15	50 - 80	70 - 150
Cost e/KW	300	300	250 - 300 (Pb) 1500 - 3000 (Li)	300 - 350	6000	400 - 1200	600 - 1500

Table 1.1: techno-economic comparison between storage technologies- Source ARER

Figure 1.8 places the various ESS according to their power or energy: small (of the order of kW, kWh) or large (of the order of MW, MWh or GW, GWh), or according to their response time: short (of the order of the second, minute) or long (of the order of the hour).

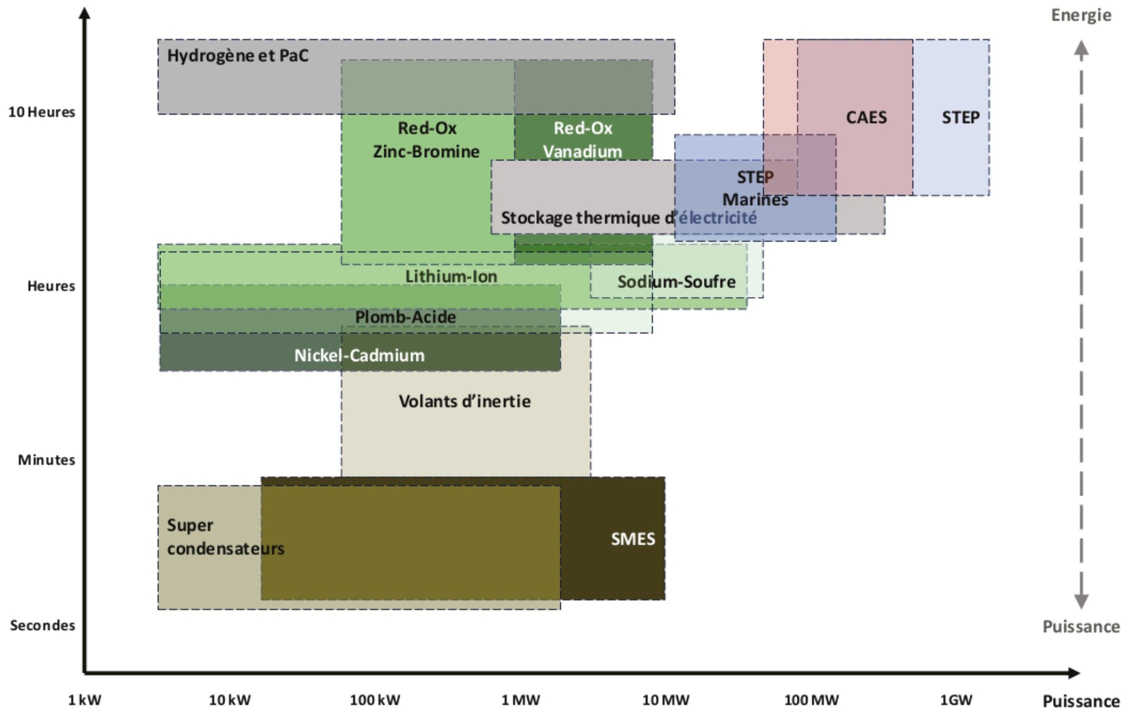


Figure 1.8: Distribution of ESS by power / response time. Source ENEA.

Energy storage is at the heart of current issues, whether it is to optimize energy resources or to promote access to them. It allows to adjust the "production" and the "consumption" of energy by limiting the losses. In fact, the energy, stored when its availability is greater than the needs, can be restored at a time when the demand is more important.

Storage methods depend on the type of energy. Fossil energy sources (coal, gas, oil), in the form of natural reservoirs, naturally fulfill the function of stocks. Once extracted, they can easily be isolated, hosted and transported from a technical point of view. Storage is more complex for intermittent energies: their production is relayed by energy vectors such as electricity, heat or hydrogen, requiring specific storage systems.

1.3/ System configuration

Two very different patterns of use must be distinguished [113] depending on the connection of power systems to the main grid, power systems can be broadly classified into:

1. Autonomous systems (stand-alone);
2. Grid-Tied systems.

Autonomous systems are designed and optimized to meet the power demand of remote places. A stand-alone systems does not have a connection to the main electricity grid (Figure 1.9). Standalone systems need to have generation and storage capacity large enough to handle the load and these systems vary widely in size. Thus, storage is necessary in order to respect the adequacy between production and consumption. In fact, Production off-cuts (partial or total according to the technology of the production system) can be created during overproduction energy times; when the power produced from the intermittent natural elements (wind, swell, solar) is greater than the power load. Conversely, load shedding may possibly be operated in case of power demand greater than the sum of the power produced (underproduction energy times).

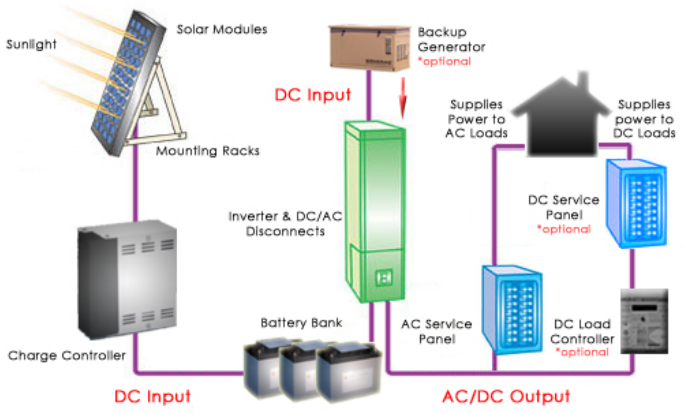


Figure 1.9: Stand-Alone or Off-Grid System [141]

In the case of **Grid-Tied systems**, the situation is significantly different because the constraints are quite different. In fact, Grid-Tied systems are connected to a larger independent grid typically the public electricity grid and feeds energy directly into the grid (Figure 1.10). In a grid connected system, the size of storage devices can be relatively smaller as the deficient power can be obtained from the grid. The feeding of electricity into the grid requires the transformation of DC into AC by a synchronizing grid-tied inverter (also called grid-interactive inverter). If electricity is supplied by "wind", "sun" or "swell", storage is useless. However, the presence of storage offers the opportunity to participate in "system services" (network regulation) and better value the energy produced economically. In this context, the sizing of the storage is a problem different from that previously described, particularly through the criteria and constraints, but the parameters to be optimized remain the same. Such a device also allows the islanding in case of disconnection of the network. In this case, we find ourselves in the autonomous situation described above.

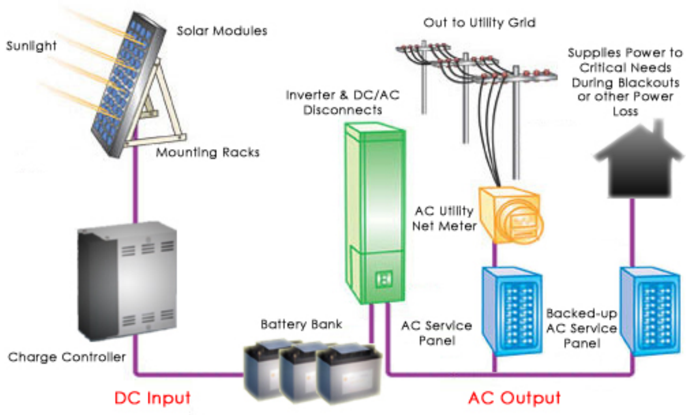


Figure 1.10: Grid Tied System [141]

1.4/ Thesis Issues and objectives

In this context, this thesis aims to design a reliable sizing of the hybrid renewable energy sources and to optimally manage it for a 100% green data center supply. This work was partly supported by the French Research Agency under the project DATAZERO.

1.4.1/ ANR DATAZERO project

The ANR DATAZERO[9] project (for DATA center with Zero Emission and RObust management using renewable energy) is building an innovative adaptive data center concept that optimizes both IT and power loads [103]. Relying on the latest technologies in terms of renewable power sources, electrical equipment, servers and cloud management systems, it aims at drastically reducing the energy consumption of tomorrow's data centers. Recently the industry has focused on increasing energy efficiency in data centers but the total amount of electricity needed for operating and hosting Cloud services is still increasing, driven by the increased size of the infrastructure to support user demands. Efforts for improving efficiency have been conducted independently in the following areas: compute infrastructure (service placement and scheduling) and renewable energy production, little has been done to focus on the interaction between these two parts. DATAZERO aims to provide new reliable solutions for high availability of IT (Information Technologies) services, andq to avoid unnecessary infrastructure redundancies while taking into account the intermittent nature of renewable power production (solar panel, wind turbine). Concerning the data center's local energy storage and management, Fuel Cells combined with Li-Ion batteries can go beyond traditional technologies, such as diesel generating sets combined with lead-acid batteries. In this way DATAZERO introduces a new concept called “data center long-term green energy storage”.

1.4.1.1/ Objective of the project

The question that we address in DATAZERO is:

How to design and manage the electricity and the IT service flows in order to deliver robust and efficient IT services to customers in data centers operating only with renewable energy sources?

To answer this question, we identified 7 scientific challenges:

- to make demand and production envelope constraints coincide with the electrical and IT planes;
- to correctly size the equipment;
- to optimally control the electrical converters to minimize electrical losses and maximize renewable energy use;
- to schedule and manage the IT load;
- to take into consideration thermal management;
- to study the complexity of the optimization problem;
- to develop a simulation toolkit.

The main targets of DATAZERO are mid-size data centers (about 1 MW) where the IT load can be managed through Virtualization or Cloud orchestrator commonly encountered in enterprise and public institutions.

[9]http://www.datazero.org

1.4.2/ Collaborators and work packages

The project is composed of 4 work packages:

- **Services / IT tasks:** The objective is the optimal distribution of computer tasks on calculators and the extraction of energy consumption. Then it will be necessary to define a power supply structure of different computers in an intelligent and redundant way.
- **Renewable & Storage Sizing:** This module aims at establishing the models of the different sources (Photovoltaic Panels, Wind Turbine, Fuel Cell) and storage means (electrolyzer, batteries, hydrogen tank). This at different time scales, from a dynamic reaction point of view (dt <1s) until their aging is taken into account. Then it will be necessary to decide their association in unit of production (a type of source + a means of storage) as well as their interconnection with the load (choice of the appropriate static converter, control law, etc.). The objective is to be able to provide statistically or periodically significant parameters in a structured way for optimization.
- **Energy management:** This module is the link between the IT part considered as a load and the electrical part (Elec). The energy management algorithm will have to allocate the energy demand to the different sources optimally. Several approaches will be studied, from a simple instant allocation to an overall optimization on a sliding window.
- **Negotiation IT-Elec:** It is in a way the conductor of the system, which according to the data of each module and also of parameters outside (weather, cycles day-night, etc.) act on the 2 previous modules in order to change the overall management strategy: reconfiguration of the distribution of calculation tasks, use or not of new sources or change of strategy of energy management.

The execution of these various works will be done in collaboration between the various partners:

- IRIT (Toulouse): The IRIT lab (Toulouse Institute of Computer Science Research) is one of the biggest lab. in computer science in France. The SEPIA team is centered on the efficient conception and usage of distributed operating systems, for clusters, grids and clouds, and is managing the Grid5000 node in Toulouse.
- LAPLACE (Toulouse): The LAPLACE (Plasma and energy Conversion lab) UMR 5213 (A+ rated), is one of the biggest laboratory in electrical engineering (topics developed: power electronics, materials, plasma and electrical systems design, control and diagnostic).
- FEMTO-ST (Besançon): FEMTO-ST Institute, UMR6174 A+ rated is one of the biggest French lab in engineering sciences (about 700 persons). Among its 7 research departments, AS2M (Automation and Micro Mechatronics Systems), ENERGIE and DISC (Informatique des Systèmes Complexes) are involved in this project.
- EATON (Grenoble): EATON is an international leader in components and electrical systems, in particular for the data center segment, with more than 100000 employees. EATON has a R&D center in Grenoble.

1.4.3/ Thesis objectives and organization

In the DATAZERO project, our task is to build the electrical part of the infrastructure. Thus, the goal of this thesis work is to design a reliable sizing of the hybrid renewable energy sources and to optimally manage it for a 100% green data center supply. To this end, this work is articulated around four main objectives:

(1) Analyze, discuss, and evaluate the validity and applicability of different model used of the renewable energy sources, of the diverse sizing and management approaches in the literature to justify our choices.

(2) Characterize and model the chosen sizing approach. This analysis will enable fixing the best configuration capable to supply the data center using different metrics.

(3) Develop an power decision module capable to negotiate with the negotiation module and to deliver different power profiles from the management approaches used.

(4) Validation of the chosen sizing with the developed power decision module.

This thesis is made as part of an AS2M and Energy department collaboration in the laboratory FEMTO-ST.

1.5/ Synthesis

In this chapter, we first presented a review of renewable intermittent energy where we explained the energetic issues and the main two criteria needed to be taken into account for our study. Then, a review of storage energy system was made; the issues and different technologies were defined and compared. This chapter ends up by detailing the DATAZERO project and explaining the scope of our contribution in it.

Chapter 2

Green Data Center & Power Supply

This chapter presents the techno-economic models, existing and retained for the study of a data center supplied by renewable energy coupled with Energy Storage System (ESS). A definition of data center structure and an overview about their huge electrical consumption is presented at first. Then, the second part of this chapter is dedicated to a state of art about the models of the chosen system composed of photovoltaic panels, wind turbines, batteries and hydrogen systems. This chapter ends up by a review on different strategies adopted by researchers in designing or managing these systems in order to supply power demand.

Contents

2.1/ Data centers

2.1.1/ Overview

The trend toward server-side computing and the exploding popularity of Internet services have created a new class of computing systems called warehouse-scale computers, or WSCs. The name is meant to call attention to the most distinguishing feature of these machines: the massive scale of their software infrastructure, data repositories, and hardware platform. In fact, Internet service may consist of many individual programs that operate together in order to implement complex end-user services such as maps, data search or email. The hardware part of warehouse-scale computers is composed of individual computing nodes including their network and storage subsystems, their conditioning and power distribution equipment, and their cooling systems. The whole set of these systems is actually a building simply referred as data centers. Thus, data centers building are quite similar to large warehouse that collocates multiple servers and communication equipment due to their physical security needs and common environmental requirements and ease of maintenance.

Yet, traditional data center host a large number of medium or small sized applications, each running on the dedicated server, decoupled and protected from other systems in the same installation, as those buildings host hardware and software of many organizational units or even different companies. In such data centers, different computer systems have a bit in common of the software, hardware or even maintenance infrastructure and do not tend to communicate with each other at all. Nevertheless, when these buildings belong to a same single organization, they use quite homogeneous hardware and system software platform, and share a common systems' management layer.

Usually, data centers are run by government agencies or big corporations. However, they are also more and more used in order to provide a fast growing cloud solution service and to meet the soaring demand for IT applications which makes these building even more hungry for power and brings large electricity bills to data center operators. Therefore, energy and power usage also holds an important role in the design of data centers because energy-related costs have become an important component of the total cost of ownership of this class of systems: the expansion of demand has increased the power consumption and in recent years power has become a significant portion of a DC's costs.

Most large data centers are built to host servers from multiple companies and can support a critical load of [10-20] MW. Very few data centers today exceed 30 MW of critical capacity. Data centers are essentially very large devices that consume electrical power and produce heat [52]. A 2008 study has estimated that data centers around the world emitted 116 million tons of carbon, a little more than what Nigeria emitted in the same year [82]. Moreover, according to the estimates of the year 2010, data centers consumed 1.5 % of the total electricity used in the world within the half is attributed to cooling increasing 300 % compared to ten years before [66].

This situation raises a global alert concerning the greenhouse gas (GHG) emissions, air pollution, social concerns and other energy security issues [26, 76]. It has led many government and researchers around the world to find a new alternative in order to reduce the energy consumption impact. These initiatives treat the subject in different ways by focusing each on a research area such as the control of the processor consumption, using dynamic voltage and frequency scaling (DVFS) or on consolidation. For instance, clustering virtual machines to avoid powering on too many physical hosts. Nevertheless, these initiatives tends to have a limited impact, since it decreases performance. It may even lead to rebound effects as users tend to use more computing resources if they are cheaper. It is thus not an adequate solution to satisfy the ever growing need for computing power while reducing the carbon emissions. Other important approaches are used to make data centers more virtuous in terms of their ecological footprint. These initiatives are explained in the following :subsections.

2.1.2/ Energy efficiency strategies

2.1.2.1/ Cooling supply strategies

A data center is not just a place of data storage, it is also a real factory of heat releases. Thus, to be able to ensure the proper functioning of the IT and electrical equipment used, industry and researchers are currently studying several methods focusing in reducing cooling demand:

- **Free cooling**: It is a methodology consisting in using natural cooling resources to decrease the heat released by the data center. Zhang et al. [145] treated the advancement of free cooling methods focusing on the configuration and performance feature. They can be divided as follows:
  - Airside free cooling: using the outside air for data center cooling can be done following two different ways. The first manner consist in aspiring the cold air outside directly in the data center (Direct airside free cooling). For instance, Siriwadrana et al. [122] considered this strategy for their data center cooling based in Australia. They investigated and analyzed the hourly temperature and humidity across all the Australian regions in order to determine the potential of using airside free cooling. They ended up by discovering a sizable potential of savings on cooling costs in some states. The second consist in operating through air to air heat exchangers (Indirect airside free cooling).
  - Waterside free cooling: it consists in using natural cold source through a cooling water infrastructure. It can be done either by direct water cooled system composed of exchangers between warm air and cold water from sea, rivers, etc. For instance a Stockholm data center reduced 80 % of energy costs while implementing a seawater cooling system. Either while using an cooling tower system allowing the refreshing of the water circulating through cracks and heat exchangers.
- **Hot and cold aisle containment**: it consist in separating hot and cold by aisle containment. It is one of the most used energy efficiency method in data center. The purpose of alternating aisle containment is to separate the source of cooling air from hot air, preventing the cold supply air and hot return air from mixing.
- **Increasing IT temperatures**: this method considers increasing the IT room ambient temperature which will directly impact the white space temperature of the data center. In fact, raising the maximum allowable white space makes it possible to use free air cooling at a higher outside temperature. In some cases, this may end up by not even installing a mechanical cooling. In [99], a study on the impact of an increased ambient temperature was made after analyzing a complete energy picture for a data center. Actually, the authors indicates the existence of an optimum temperature depending on individual characteristics, including IT equipment, cooling system architecture and location.

To summarize, these methods has led to significant power and cost savings on data centers cooling demand.

2.1.2.2/ Power supply advanced concept

In order to minimize the electrical consumption of data centers, some researches worked on the advancement of electrical devices of the data center or in the power distribution:

- **Smart management using the Uninterruptible power supply (UPS)**: The UPS is an electrical device generally used as an emergency power in case of fail-over situations such as diesel generators. Nevertheless, the strategy here consists in using the UPS in a different manner as an energy storage device. For instance, in the article [128], an investigation about cost reduction opportunities using the UPS unit as energy storage device was established. The authors used Lyapunov optimization to develop an on-line control algorithm able to minimize the time average cost. Moreover, Deng et al. [28], tried to decrease the operation cost of data center by introducing an online control policies that uses several characteristics of several energy sources. The authors use Lyapunov optimization, and design a control algorithm that does not require significant statistics of the dynamics of the system. The control algorithm makes decisions on *long-term ahead power procurement*, *real-time power procurement* and *real-time uninterrupted power supply (UPS) charging/discharging*. A one-month trace-driven simulation shows that the proposed algorithms provide a compromise between operational cost, UPS lifetime and data center availability. The algorithms are also robust to time-varying supply and power demand.
- **Direct current optimization**: The power distribution of data center can be made using alternating current (AC) or direct current (DC). Although, the authors in [111] believed that AC displaces the DC and will stay the best method of powering a data center thanks to its compatibility. In this context, Tom et al. [126] proposed in 2008 a comparison of two different power installations. In fact, in the first installation, direct current was distributed at the facility level to racks of computers which accept high voltage direct current. In the second installation, the DC power power conversion occurs at the rack level and is then distributed to the server of the rack. They showed that when DC is then directly distributed to servers within the rack can create electrical inefficiency through the converters AC/DC-DC/AC.

To end up, these methods has led to significant power and cost savings on data centers power supply.

2.1.2.3/ Managing workloads in data centers

Recently, numerous works in managing IT resources are being done in order to optimize the scheduling computation or allocation such as:

- **Scheduling computation tasks under energetic constraints**: In fact, in [61], several common standard scheduling techniques such as list scheduling algorithm and binary search technique were adapted to take power constraints into account. They were tested on a high performance multi-core computing platform that is powered solely by renewable sources. Their approach does not aim to reduce the energy consumption, but rather to optimize typical common scheduling objectives such as minimizing the makespan and the total flow time, while respecting the power constraint. The amount of jobs that can be executed simultaneously at a certain time is limited by how much power is being produced by the renewable sources, which naturally varies through time. Their results show that the list scheduling algorithms that they adapted to this problem give fast and good solutions. In [62], Kassab et al. used Genetic Algorithm (GA) to deal with the IT jobs scheduling problem in green data centers. The paper assumes that different tasks have different power consumptions, according to their computational density; therefore, changing the order of the tasks leads to different power consumption levels at a given time. In the proposed GA, each chromosome represents a different solution. Each solution proposes a possible order of the tasks of a set of independent sequential tasks. The GA then shuffles the genes of a chromosome (the order of tasks) by applying several different crossover and mutation operators that are tested and compared in this work. Then, the tasks are greedily scheduled according to their order in each chromosome on a multi-core platform powered entirely by renewable energy sources.

- **Problems of allocation and resource management** (server shutdown, migration of virtual machines, etc.): In [46, 47], the authors propose GreenSlot which is a batch scheduler for parallel tasks. The aim is to reduce the brown power consumption of a data center partially powered by solar panels. In GreenSlot computation jobs have deadlines and the scheduler first reserves resources for these jobs with lower slack (distance from latest possible start time to current time). Based on weather forecasting and power prediction, GreenSlot schedules the tasks on time slots. However the authors do not try to optimize their schedules. They just reduce the consumption and costs while meeting as much deadlines as possible. The simulation results are generated for two different systems: SLURM (scheduler for Linux) and MapReduce (scheduler for Hadoop). Similarly, Liu et al. in [77] present an holistic approach to optimize the energy cost with incomes from running a batch of jobs and outcomes to buy brown energy.

To conclude, these methods has led to significant power and cost savings on data centers supply.

2.1.2.4/ Managing en energy sources in data centers

Two approaches are usually followed when trying to reduce the use of brown energy in data centers, besides compute/load management: using local renewable resources (on-site generation) and/or resources available on the grid (off-site generation).

- **Off-site generation**:

The choice of IT companies depends on various factors such as financial interests, grid reliability, and geographical location. For example, due to the intermittent character of Renewable Energy System (RES), the uncertainty issue can be dealt with by using additional resources and/or curtailment. However, IT companies, instead of dealing with the issue themselves, might prefer to use already existing RES on the electricity grid. This option can influence the local producers and/or regional governance to build more renewable resources to attract IT companies through long-term agreements. For example, MidAmerican Energy, a utility in Iowa, decided to invest $1.9 billion to increase its wind generation after a renewable energy policy announcement of Facebook [35]. Similarly, Google[1] operates its data centers using 100% renewable energy since 2017, but this energy is not generated on-site. There is however no guarantee that the electrons supplying the data center are actually “green”, as they will typically originate from the nearest plants, green or not.

- **On-site generation**:

On the other hand, if the current electricity grid is not able to support a high penetration of RES, if the data center operator requires the use of clean energy or if the operator requires greater control on energy supply, companies might be interested in investing in on-site generation, either from conventional or from renewable generators. Then, IT companies avoid the incurred transmission and distribution costs, as well as reduce the system losses [136]. For example, large companies such as Apple, Google or eBay are building unplugged data centers to have a better control over their infrastructure. eBay’s Utah data center uses natural gas to produce its electricity [2], while the grid is only used as a backup. In a similar fashion, Apple’s Maiden, North Carolina Datacenter[3], is using a 40 MW solar plant along with biogas fuel cells. Similarly, Apple has invested

[1] https://sustainability.google/projects/announcement-100/
[2] https://www.ebayinc.com/stories/news/new-utah-data-center-advances-our-commitment-greener-commerce
[3] https://www.datacenterknowledge.com/data-center-faqs/apple-data-center-faq-part-2

in several on-site installations of renewable energy, particularly solar power in North Carolina and Nevada [58]. McGraw-Hill has recently completed a 14 MW solar field for its data center [89]. A drawback is that this type of installation requires large financial investment, which might not be possible for some small IT companies. Furthermore, diverse researchers explore in the same direction trying to combine renewable energy sources in order to provide electricity to their data centers. The European project named “GreenDataNet” coupling solar panels and batteries is a good illustrating example. They designed and validated a new smart energy management system allowing urban data centers to radically improve their energy performance [48]. Also, the “Renew it” project developed a simulation tool to evaluate the energy performance of different technical solutions integrating diverse renewable energy sources as wind turbines and photovoltaics in different European climate regions [112].

Knowing that the European Technology Plate-form (ETP) for electricity networks of the future, more popular as ETP Smart grid expected that by 2020 approximately 34% of the total electrical consumption will come from renewable energy and will have gone more than that by 2035 [130], these initiative allows data center either to generate their own renewable energy or to use power directly from the nearest renewable power plant decreasing then their influences on climate change.

Designing green data center implies to change the electrical generators of the latter and to use other reliable technologies, environmentally friendly, cost-effective, and affordable even to the general public infrastructure. First, for a better understanding of the data center buildings, let’s detail their infrastructure.

2.1.3/ Green data center infrastructure

Data centers can be classified following different parameters such as their localization, size and system type. Redundancy represents a main issue for data centers as they have to guarantee a quality of service to their clients (if one component breaks down, the system has to continue working). Moreover, while connecting many energy providers to the data center, redundancy qualify the latter function and categorize it into four architectures [107, 118]:

- **Redundancy level 0**: N architecture 2.1 with no redundancy or backup. Only the number of components needed is provided, there is no resilience and component failure will cause downtime. As there is no duplication of any of the components, the system will stop working when a component break down (see Figure 2.1).
- **Redundancy level 1**: N+1 redundancy means that the number of components provided is one more than the number of components needed so it is a N+1 structure. This kind of system has a low level of redundancy. It could still work when a component who has the “+1” goes into a breakdown (see Figure 2.2).
- **redundancy level 2**: 2N redundancy means that the data center architecture consists of two complete power systems, each containing N components, run in parallel to hot swap between each other in case of failure. (see Figure 2.3).

According to D. O’Sullivan in [97], most data centers use either N+1 or 2N architectures. As stated in [98], redundancy is necessary to achieve higher availability, but it also improves the maintainability, flexibility, expandability and fault management of data centers. However, it is also a complex trade-off due to its impact on Total Cost of Ownership. In his report [110] written in 2011, Neil Rasmussen describes the mean load of a data center, which is usually far from 100%; consequently, having a 2N redundancy leads to an increase in infrastructure that is already usually oversized.

New architectures are proposed to change the classical paradigm of having a grid-connected data center using fuel generators for backup. As further described in Section 2.1.2.4, renewable energy can supplement or replace grid supply and backup generators. This energy can also be generated and used locally or imported from external sources.

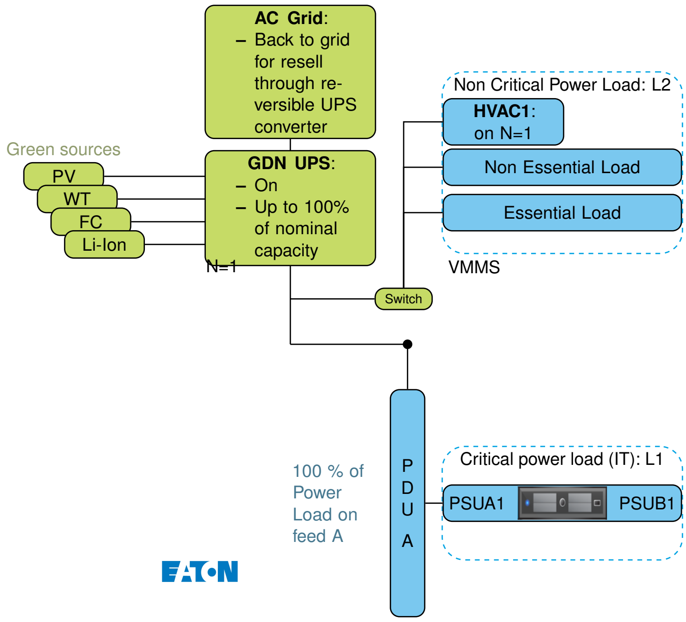


Figure 2.1: Classical N architecture.

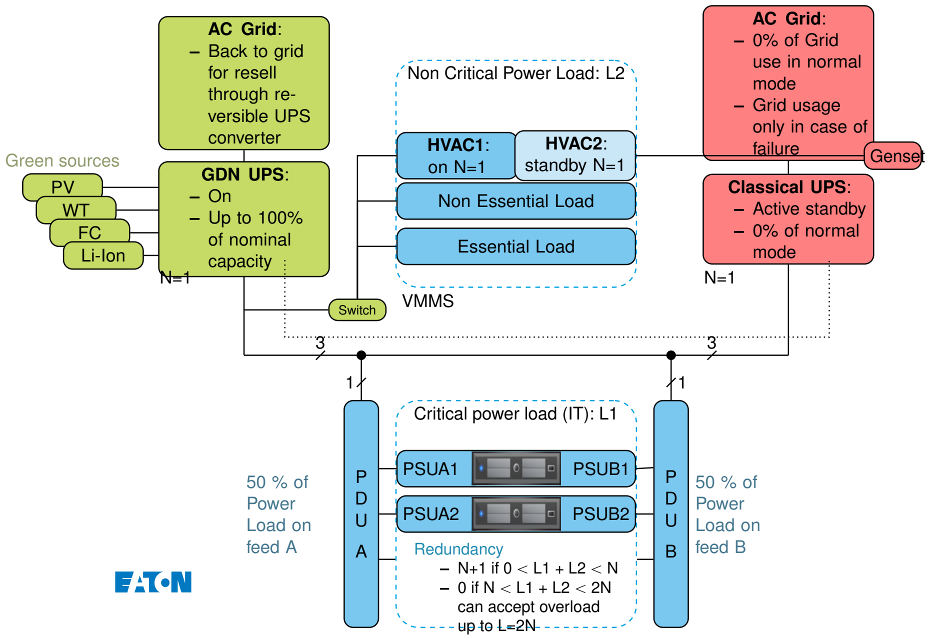


Figure 2.2: N+1 architecture: The Critical load can be fully supported by green sources (left part) which can be replaced in case of failure by a fail safe system (right part) such as a fuel-based or H2-based generator.

In addition to these considerations, installing data centers in remote places such as a research center in Arctic or in countries after major devastating events leads to face multiple challenges. The concept of containerized data centers is often used in these contexts. All the computing, cooling, and electrical infrastructure of a small data center is included in the form-factor of a transport container [131, 133]. For example, using wind-powered containers, in the patent [16], Burger et al. have proposed in 2011 a self-sufficient container storing the energy surplus in batteries. Other types of containers are also leveraging their form factor to achieve free-cooling such as the underwater container built by Microsoft [83], which still needs to be plugged to the electrical grid.

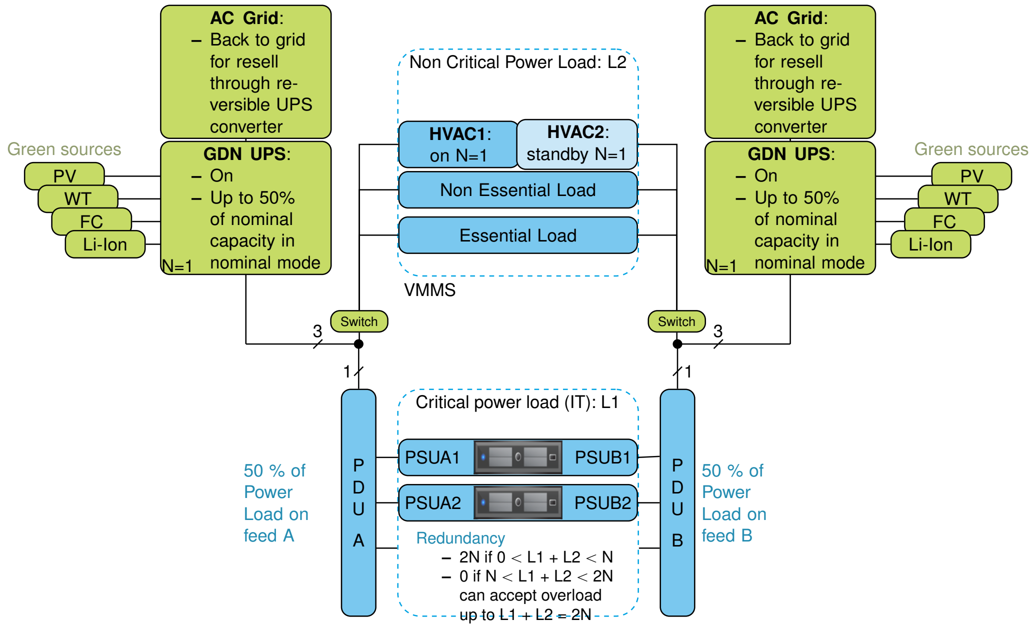


Figure 2.3: 2N architecture: The critical load can be fully supported by each of the green sources on right and left. Even in case of failure of half of the sources, the data center can still operate using 100% green energy.

More extreme concepts have also been explored. For example, in [20, 100], authors propose to use a swarm of satellites carrying computing nodes. Using this technology helps having a completely deterministic renewable energy production. This approach is however limited due to the constraints of building and maintaining a swarm of satellites.

2.1.4/ Data center selected model for DATAZERO

In the ANR DATAZERO[4] research project, we intend designing a data center with the following constraints:

- 1 MW total electric capacity distributed on four (2x 2x 275 kW in peaks) branches 2N redundancy.
- Each N must provide 500 kW and will be exactly duplicated.
- The cooling is taken into account into these 500 kW.
- The data center is only supplied by renewable energy produced locally.
- The chosen renewable energy sources are solar panels and wind turbines.
- The chosen storage system sources are hydrogen system composed of fuel cells, electrolyzers and hydrogen tanks and battery system (lithium-Ion).

The following sections intends to propose power and energy source models (profiling) meeting DATAZERO project needs.

[4] http://www.datazero.org

2.2/ Hybrid renewable energy system models

In DATAZERO project, the hybrid renewable energy system consists of photovoltaic panels (PV), wind turbines (WT), hydrogen storage systems (HSS) and battery storage systems (BSS) as shown in Figure 2.4.

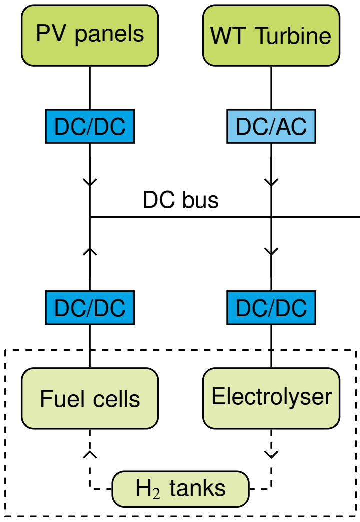


Figure 2.4: Structure of the DC microgrid supplying the datacenter loads.

Then, this section introduces the models of the chosen hybrid renewable energy system from literature data. Characteristic equations are thus established for each component, from renewable sources to storage units for two different purposes. The first purposes consist in **sizing and long term management** written (A) in the modeling tables. The mathematical models presented in here do not take into account the instantly variation. The models are on the hourly, daily scale in order as the users apply it for a sizing methods and/or long term management of an infrastructure. The second purposes consist in the **control of the devices and the short term management** written (B) in the modeling tables. These mathematical models should be really detailed as the times scale here is micro or milliseconds. The control and/or the management needs to be done for the next hour.

Both purposes (A) and (B) are defined and explained in each power sources modeling in order to choose, at the end, the most suitable model. These equations are expected to be used in optimization loops for energy management and sizing.

2.2.1/ Wind power model

The local climate conditions have to be taken into account in the used models as the data center can be installed in diverse location. Actually, wind speed differs through the pressure and temperature of each chosen location. Moreover, wind speed measurements are obtained at a 10 m height, which mismatch with turbine hub height. In these cases, The wind speed $v(t)$ can be estimated as follows:

$$v(t) = v_0(t)\frac{log(\frac{h}{z})}{log(\frac{h_0}{z})} \tag{2.1}$$

where v_0 is the measured or forecasted wind speed, in m/s, h is the desired hub height, in m, h_0 is the actual height of measurement, in m and z is the terrain roughness length, in m.

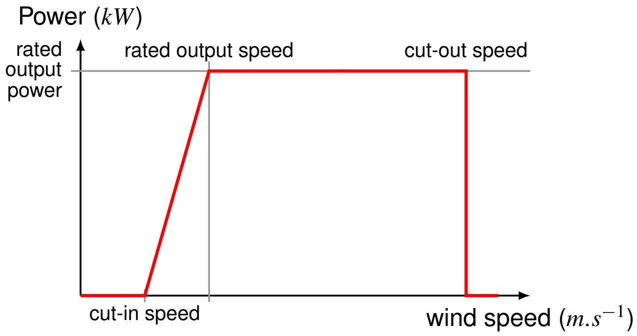


Figure 2.5: Ideal wind turbine power output

After environmental conditions and input data are known, the turbine model can be established. Although they are very similar, several basic models are used in the literature. In the following, we define two models in the table 2.1: a simple one, suitable for sizing and long term scheduling (A) [31, 45, 125, 143], and another one, more detailed and suitable for precise control of equipment (B) [56, 57, 84].

Table 2.1: Wind turbine mathematical models following the use purpose

(A)	(B)
$Pwt_k = \begin{cases} 0 & \text{if} \quad v(k) \leqslant v_{ci} \text{ or } v(k) \geqslant v_{co} \\ P_r \dfrac{v(k) - v_{ci}}{v_r - v_{ci}} & \text{if} \quad v_{ci} < v(k) < v_r \\ P_r & \text{if} \quad v_r < v(k) < v_{co} \end{cases}$	$Pwt_k = \frac{1}{2}\rho\pi r^2 c_p v^3(k) \times 10^{-3}$
v, the wind speed value, in m/s. P_r, the rated output power of the WT, in kW, v_{co}, the WT cut-out wind speed, in m/s. v_{ci}, the WT cut-in wind// speed, in m/s, v_r, the WT rated wind speed, in m/s,	ρ, the air density, in kg/m 3 v, the wind speed value, in m/s. r, the length of the WT blades, in m c_p, the conversion efficiency.

Both models take the wind speed as input and deliver the output power of the wind turbine (WT). In fact, the output power model of one wind turbine generator follows the power curve shown in Figure 2.5. So, the turbine starts generating power at the "cut-in" wind speed v_{ci}. Then, the generated output power increases with the increase of wind speed from the "cut-in" v_{ci} to the rated wind speed v_r. When the wind speed varies between the rated wind and the "cut-out" wind speed v_{co}, which is the maximum wind speed value at which the turbine can correctly work, the turbine produces a constant or "rated power". Once the wind speed goes beyond the "cut-out" speed, the turbine stops generating for safety reasons.

2.2.2/ Photovoltaic panels model

A photovoltaic panel converts solar energy into electricity. Multiple panels can be connected in series and/or parallel to form a plant. In the following, two models are proposed: the simple one (A), used by many researchers in their sizing or management work such as [10, 60, 115, 121]. The second one (B) more detailed was used by many researchers [79, 127, 129] for precise control of equipment These two models are summerized in the Table 2.2

Table 2.2: Photovoltaic mathematical models following the use purpose

(A)	(B)
$Ppv_k = I_k \times A_{pv} \times \eta_{pv}$	$Ppv_k = V_{pv}(k) \times I_{pv}(k)$ with: $I_{pv}(k) = N_p I_{ph}(k) - N_p I_s(k)(exp(\frac{qV_{pv}(k)}{N_s AkT_c(k)}) - 1)$ $I_{ph}(k) = I_{sun}(k)(I_{sc} + K_i(T_c(k) - T_{ref}))$ $I_s(k) = \frac{I_{sc}}{exp(\frac{qV_{pv,oc}}{N_s kAT_c(k)})-1}(\frac{T_c(k)}{T_{ref}})^3 exp(\frac{qE_g}{kA}(\frac{1}{T_{ref}} - \frac{1}{T_c(k)}))$
I, is the solar radiation in W/m^2, A_{pv}, is the area of the PV panels in m^2, η_{pv}, is PV panels efficiency,	I_{sun}, the forecasted solar radiation, in W/m 2 T_c , the panels cells temperature, in K V_{pv} , the plant voltage, in V N_s , the number of the cells in series, N_p , the number of cells in parallel, T_{ref} , the reference cell temperature (303.15 K), I_{sc} , the short circuit current, in A, $V_{pv,oc}$, the open-circuit voltage, in V, A, the ideal factor, E_g , the band gap energy of the semiconductor, in eV, K_i , the temperature of the circuit current, in A/K, k, the Boltzmann constant, in J/K, q, the elementary electronic charge, in C, Ppv , the plant power output, in kW, I_{pv}, the plant current, in A.

2.2.3/ Hydrogen storage system model

Component models of the hydrogen storage system are explained a follows:

2.2.3.1/ Fuel cell model

Table 2.3 summarizes a simple model (A), suitable for sizing and managing [24], and another one suitable for controlling electrical device [69] is also introduced (B).

Table 2.3: Fuel cell mathematical models following the use purpose

(A)	(B)
$Pfc_k = \eta_{fc} \times LHV_{H_2} \times Qfc_k$ $Pfcmin < Pfc_k < Pfcmax$	$Pfcmin < Pfc_k < Pfcmax$ $Pfc_k = \eta_{sys} \times V_{fc}(k) I_{fc}(k)$
η_{fc} , the fuel cell system efficiency, LHV_{H_2} , the lower heating value of hydrogen, in kWh/kg (constant at 33.33kWh/kg), $Pfcmin$, $Pfcmax$, the fuel cell system min. and max. power output, in kW, Qfc, the hydrogen consumption, in kg	I_{fc} , the fuel cell current, in A, V_{fc} , the fuel cell voltage, in V, η_{sys} , the system efficiency $Pfcmin$, $Pfcmax$, the fuel cell system min. and max. power output, in kW

2.2.3.2/ Electrolyser model

In the following table 2.4, a simple electrolyzer system model suitable for sizing and managing is given on colomn (A) [8,24,81,93], and another one suitable model for the control part (B).

Table 2.4: Electrolyser mathematical models following the use purpose

(A)	(B)
$Pez_k = \frac{\eta_{fc} \times Qez_k}{HHV_{H_2}}$ $Pezmin < Pez_k < Pezmax$	$Pezmin < Pez_k < Pezmax$ $Pez_k = V_{el}(k) I_{el}(k)$
η_{el} , the electrolyzer system efficiency, HHV_{H_2} , the higher heating value of hydrogen, in kWh/kg (constant at 39.4kWh/kg), $Pezmin, Pezmax$, the system min. and max. power output, in kW, Qez, the hydrogen production, in kg	$Pezmin, Pezmax$, the system min. and max. power output, in kW, I_{el} , the current density, in A/cm 2 V_{el} , the output cell voltage of the electrolyzer, in V.

According to the studies mentioned above [24,81], the electrolyzer has physical constraints that have to be taken into account during the operation; the electrolyzers have not been designed to operate under variable power conditions, such as those that can be found in renewable micro grids. For safe operation, the electrolyzer should be operated above a minimum power level in order to avoid impurities in the gases produced and hazardous mixtures of H_2 in O_2. In addition, high current densities accelerate the degradation due to over-voltage in the electrodes.

2.2.3.3/ Hydrogen level of H_2-tank

The hydrogen produced by the electrolyzer and consumed by the fuel cell is stored in tanks. For sizing and scheduling application purposes (A) [18,73] and for short term control and simulation applications the model of hydrogen tanks is defined in equation (2.2) where the hydrogen level LOH-tank at time step t depends on hydrogen level at time $t-1$ ($LOH(t-1)$), output hydrogen mass flow of electrolyzer at time step t (Q_{ez}), and hydrogen consumption of FC at time t. (Q_{fc}):

$$LOH_k = LOH_{k-1} + Qez_k - Qfc_k \tag{2.2}$$

2.2.4/ Battery storage system Models

Several researchers [10,49,72,115,143] used the simple model (A). In this model, the status of the battery is represented by its State-of-Charge (SOC), which represents how full it is. The SOC varies with how electric power is input or drawn into/from the battery, while considering the related efficiencies. Note that the starting time is negligible and the ramp rates can be assumed infinite. For short term control and simulation applications (B) [6,39,121,139], the voltage of a battery depends on several parameters, such as the state-of-charge (SOC, which represents how charged or discharged the battery is), the current, and the temperature.

These both models are defined in Table 2.5.

2.2.5/ Selected models for DATAZERO project

Hybrid renewable energy sources is largely dependent on its components, thus an accurate modeling of each component of HRES provides tools to better understand the performance and reliability of the system, assisting to optimize it. Thus, after analyzing the models and following the

Table 2.5: BSS mathematical models following the use purpose

(A)	(B)
$SOC_k = SOC_{k-1} \times (1-\sigma) + \frac{Pch_{k-1} \times \eta_{ch} \times \Delta t - \frac{Pdch_{k-1}}{\eta_{dch}} \times \Delta t}{Cbat}$ $SOCmin < SOC_k < SOCmax$ $SOC_0 = SOCinit$	$SOC_k = SOC_{k-1} + \eta_{ch/dch} \frac{Ibat_k \Delta t}{Cbat}$ $SOCmin < SOC_k < SOCmax$ $SOC_0 = SOCinit$ $I_{bat,min} < Ibat_k < I_{bat,max}$
Pch , the charging power (<0), in W, $Pdch$, the discharging power (>0), in W, $Cbat$, the battery capacity, in Wh, Δt, the time step duration, η_{ch}, the charging efficiency, η_{dch}, the discharging efficiency, E_{bat} , the battery capacity, in Wh, $SOCmin$, $SOCmax$,the min. and max. SOC values, in %,	$Ibat$, the battery current, in A. $Cbat$, the battery capacity, in Wh, $\eta_{ch/dch}$, the charging/discharging efficiency, $SOCmin$, $SOCmax$, the min. and max. SOC values, in %, $I_{bat,min}$, $I_{bat,max}$, the battery min. and max. current, in A.

thesis objectives (Optimization of sizing and management), the models chosen for each sources are the simple model (A) defined for sizing and scheduling applications purposes. Thus, in the following sections, a review about performance metrics used in the literature to evaluate the reliability and/or feasibility of the HRES, sizing methodology and optimal management are presented.

2.3/ Electrical sizing

2.3.1/ Definition

Sizing a hybrid renewable energy system consists in finding the most appropriate size, taking into account the intermittent production and forecasts, that allow to maximize (or minimize) predefined performance criteria. Sizing of HRES, whether in grid connected or in stand alone system follows a common methodology used by the scientific community. Three types of input parameters are distinguished [90]:

- Economic parameters (performance metric);
- Meteorological parameters (measurements and/or forecast);
- Production parameters (non-renewable, RES, etc.).

An optimization procedure is then performed to determine the best configuration based on predefined criteria. The output parameters are the platform cost, the number of elementary modules (PV panels, wind turbines, batteries, etc.) or the distribution of production (energy/power) in the case of a distributed system, the size of the storage as well as that its associated operation over the period, as shown in Figure 2.6.

2.3.2/ Technical performance metric

There are different criteria for evaluating the performance of a HRES which can be used to identify the reliability and/or feasibility of the system. Then, the designer is able to size in an accurate way the platform. The most known criteria are briefly described in this section.

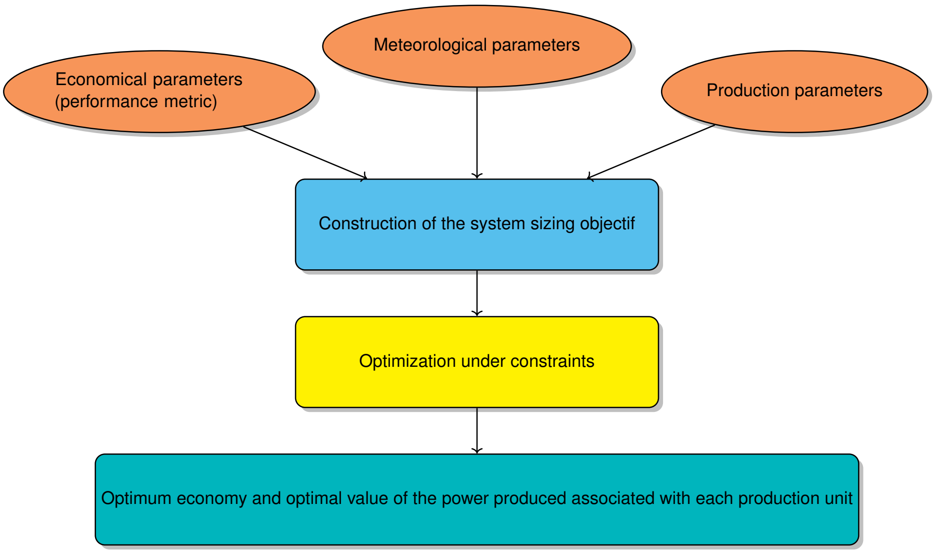


Figure 2.6: Standard methodology of sizing

2.3.2.1/ Loss of power supply probability (LPSP)

Loss of power supply probability (LPSP) criterion was introduced in [139] for isolated systems in 2003. It assesses the reliability of demand satisfaction in the case of an isolated system and represents the share of energy not provided on the requested energy [80]:

$$LPSP = \frac{\sum_{k=0}^{K-1} Pload_k - Pprod_k}{Pload_k} \text{ (s.t. } Pload_k > Pprod_k) \tag{2.3}$$

where the $Pload_k$ is the energy demand and $Pprod_k$ is the energy produced by the HRES. Another definition of LPSP appeared in [138] and represents the time of service default. It is defined as follows:

$$LPSP = \frac{\sum_{t=0}^{K-1} \Delta k}{\mathcal{H}} \text{ (s.t. } Pprod_k < Pload_k) \tag{2.4}$$

where Δk is the time step and $\mathcal{H}$ is the horizon considered. A LPSP equals to 0 means that the system is absolutely reliable and a LPSP equals to 1 means that the demand is not satisfied at all by the generated power.

2.3.2.2/ Loss Of Load Probability (LOLP)

The Loss Of Load Probability (LOLP) is the probability that a loss of load situation will occur, i.e., a time when the consumption is greater than all the means of production available, which may lead to a load shedding (inability to meet demand). This criterion is also used in the security of the national electrical systems for which the number of admissible hours of loss of load over a year (Loss Of Load Expectation or LOLE) is 3 hours (LOLP = 0.034 %).

2.3.2.3/ Level of Autonomy (LA)

Level of Autonomy (LA) is defined as the complement to 1 of the ratio between the number of hours of Loss of Load divided by the total number of hours the facility is operational:

$$LA = 1 - \frac{H_{LOL}}{\mathcal{H}} \tag{2.5}$$

2.3.2.4/ Annual system cost

The annualized cost of the system (ACS) includes the capital cost (C_{cap}), the replacement cost in case of system breakdowns (C_{rep}) and operation and maintenance costs to ensure the lifetime of the system ($C_{O\&M}$) defined by:

$$ACS = C_{cap} + C_{rep} + C_{O\&M} \tag{2.6}$$

In the design of an electrical system, evaluating criteria are essential for security of supply, reliability of the system, etc. In an isolated mode (SA) but also connected (GC), it is a question of a best answering to a demand of final consumers. In the following, we studied different way of sizing an HRES.

2.3.3/ Existing methods methodology

Diverse methodologies have been used in order to compose optimal energy systems. An easy way to classify them is according to their complexity level as we can see in Figure 2.7.

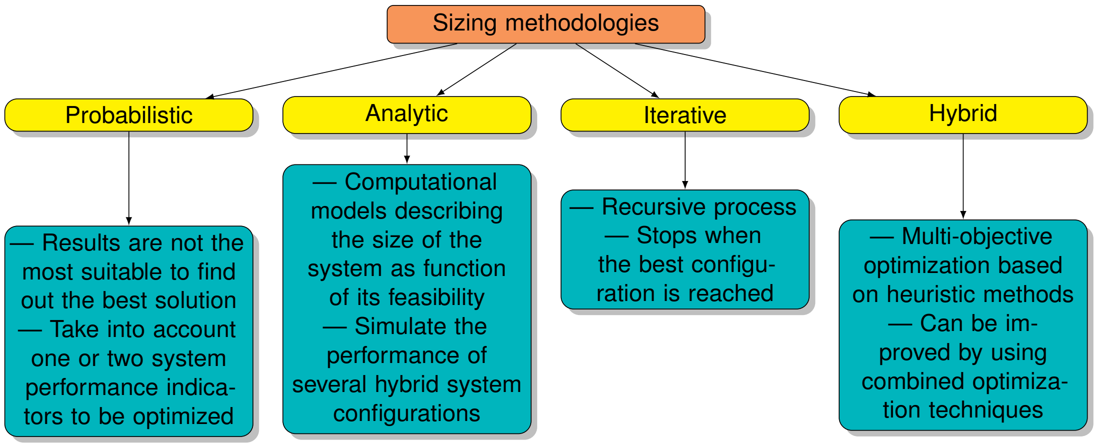


Figure 2.7: Sizing methodologies

2.3.3.1/ Probabilistic methods

A probabilistic method is an efficient technique based on probability to prove the existence of combinatorial objects having some specific properties. This method consists in generating entire sample space for combinatorial object, places probability on each object in a simple space. Then given these probabilities, we verify if the probability of the combinatorial object exists or if we consider complement. In sizing of HRES, these methods can be one of the simplest sizing method. Nevertheless, the obtained results are not efficient to find the best solution. Generally, these methods take only one or two performance criteria to be optimized into account.

For instance, Yang et al. [139] proposed a probabilistic method, where they prove the importance of choosing a suitable typical meteorological year (TMY) in order to get an accurate assessment of performance in a hybrid PV-wind energy system. Another probabilistic approach is suggested by Tina et al. in 2011 [125]. Their method is based on convolution techniques using probability density functions (PDF) to assess long term performance of hybrid solar wind power systems.

In addition, C. Protogeropoulos et al. presented two sizing methods [106] for autonomous PV-wind hybrid systems many years before. Both methods determine system performance by computing the system autonomy level and the total cost. However, in the first method, an annual average of monthly system performance is used and, in the second one, the worst monthly scenario.

2.3.3.2/ Analytical methods

The performance evaluation of a hybrid system in this case can be achieved by computational models (i.e., commercial software tools and/or numerical approximations of system components) describing the HRES size in function of its feasibility. Therefore, system performance can be evaluated for a conceivable system architecture set and/or an appropriate component size. The best results are obtained due to the analyze of one or several performance criteria.

Recently, several computer tools have been developed in order to help the user to analyze the integration of the sources for optimizing, designing and evaluating performance of PV-wind hybrid systems as discussed in a comprehensive review by Sinha and Chandel [120]. The most used ones are: Hybrid Optimization Model for Electric Renewable Energy (HOMER) [96], software developed by the National Renewable Energy Laboratory (NREL), The Hybrid Power System Simulation Model (HYBRID2), Hybrid Optimization using Genetic Algorithm (HOGA), etc. However, these softwares have limitations like black box coding, different working platforms, unavailability of some of the software, and they are also not as flexible as optimization techniques which can be used as per research criteria [120].

2.3.3.3/ Iterative methods

Iterative methods in sizing of HRES consist in using recursive process taking design specifications into account that stops once the best configuration is achieved.

Many researchers have worked using genetic algorithms (GA) to achieve a good sizing as close the optimal solution as possible. For instance, Kaldellis et al. [60] proposed to minimize the system cost by means of electrical load under some design constraints. Their best sizing solution is reached by using a GA process. Similar works can be found in Dufo-Lopez and Bernal-Agustin in 2005 [34] and in Yang et al. [137, 138].

Moreover, Ashok [4] obtained a hybrid system among different combinations for a rural community, minimizing the total life cycle cost and ensuring systems reliability. In this work, a numerical algorithm based on the QuasiNewton method was used to solve the optimization problem [109].

Another iterative method is reported in [75]. The authors analyze real-world workload traces from a medium-sized data center to calculate the optimal size of the solar panels needed to generate enough green energy to cover the entire workload energy consumption, assuming that all the surplus solar energy is stored in a battery of infinite size. Then, assuming that the size of the used solar panel is the found ideal size, they calculate the optimal capacity of the battery needed so that the workload consumes zero brown energy. They show that using an opportunistic scheduling algorithm demands a smaller battery size than a baseline algorithm with Energy Storage Device (ESD). Finally they combine both approaches to find the optimal solution for solar panels and different batteries of different sizes.

2.3.3.4/ Hybrid methods

Using hybrid method in HRES is defined as an optimization problem with a multidimensional nature. It consists in solving a multi-objective optimization based on combining several methods such as heuristic method (GA, neural network, etc.) and other optimization techniques.

For instance, many researchers [9, 63, 117] modified Genetic Algorithms (GA)) in order to give the designer the choice of the optimal configuration. This was done by a set of non-dominating Pareto set using selection criteria. In [63], the optimization objective was twofold and consists of minimizing the system cost of energy and greenhouse gas (GHG) emissions by six different constraints. In this work, the main originality comes from the assessment of GHG emissions based on life cycle analysis. A similar work was proposed in [134] where the set of non-dominated Pareto solutions was obtained using a particle swarm optimization algorithm (PSO). The optimization objective in this work was to twofold techno-economical and threefold technical, economical and environmental.

Moreover, Artificial Neural Networks (ANNs) can be used as a hybrid method for sizing, as it is considered as a simplified mathematical model of brain-like systems. Different authors used the ANN for many purposes. For example, [54] and [87] used ANN in the prediction of daily solar radiation in isolated areas where data were missing. It is also used to improve the analytic method of loss of load probability (LOLP) [55, 86, 88].

2.4/ Management of green energy sources

2.4.1/ Definition

Nowadays, energy management became a key issue for companies as they are playing an essential role in giving their organization more sustainable energy choices. Managing a hybrid renewable energy system consists in finding the right balance between the power demand of the clients and the power production. A good management strategy must increase profitability to its maker while respecting some defined constraints such as decreasing the brown energy consumption, reducing costs, decreasing the companies' environmental impact, etc. Thus, these strategies are established with clear targets and provide a logical framework for utilizing the appropriate resource in the appropriate time to achieve these goals.

Due to the global warming alert, a wide-range of different actors and stakeholders developed an interest in green renewable energy use and management, such as industries, clients, government authorities, researchers, NGOs, associations, local communities, suppliers, etc. Therefore, many management strategies have seen the light and are considered, used and improved every day to ensure the reliability and sustainability of the HRES. In these management strategies, many optimization methods were used in order to get results as optimal as possible. Figure 2.8 summarizes some of the optimization techniques used in the research community with different approach in order to improve the commitment of the renewable energy sources.

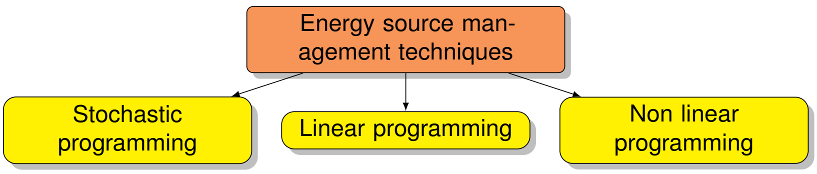


Figure 2.8: HRES management approach

2.4.2/ Existing management methods

2.4.2.1/ Stochastic programming

Stochastic programming is an algorithmic approach that incorporates uncertainty by nature. The uncertainty is unknown and usually characterized by a probability distribution on the parameters. Authors used stochastic program to minimize the exposure of each microgrid to vulnerabilities. For instance, a stochastic energy and reserve scheduling method for a microgrid, which considers various types of demand response algorithms (DR), has been proposed in [142]. In this proposed scheme, all types of customers namely residential, commercial and industrial ones can take part in demand response programs which will be considered in either energy or reserve scheduling. A probabilistic energy management system, to optimize the operation of the microgrid, based on an efficient point estimate method is proposed in [65, 91]. The authors used the proposed method to model the uncertainty associated with the wind, solar power generation systems, the market prices and the load demands. Moreover, in [2], the authors have proposed a strategy for constructing microgrids with optimized supply adequacy. The problem has been formulated as multi-objective optimization problem. The authors used three different types of algorithms at different stages, graph theory related algorithms and forward-back-ward-based probabilistic algorithm. They argued that the proposed planning framework can help utility engineers and system planners in designing microgrids capable of working in island mode.

2.4.2.2/ Linear programming

Linear programming is a mathematical problem of optimizing (maximizing or minimizing) a linear function of several variables that are connected by linear relationships called constraints. For example, the two-variable problem $x = (x1, x2) \in R^2$ such that:

$$\begin{cases} \text{minimize} & x_1 + x_2 \\ s.t.: & \\ x_1 + 2x_2 \geqslant 2 & \\ x_1 \geqslant 0 & \\ x_2 \geqslant 0 & \end{cases} \tag{2.7}$$

which consists in minimizing the linear function $x = (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1 + x_2 \in \mathbb{R}$ under the linear inequality constraint $x_1 + 2x_2 \geqslant 2$ and the positivity constraints of x_i. It is in fact a linear optimization problem. A polynomial time algorithm is able to optimally solve this problem as long as the variables are real-numbers. As it is a convex problem, the solution is a point of the convex polytope defined by the constraints. It is worth $(x1, x2) = (0, 1)$. On the other hand, as soon as the variables are integer or boolean, the problem becomes more difficult (NP-Complet) and it is then only possible to find an optimal solution in cases where the number of variables is small. Indeed there is no reason that the points of the polytope should be integer.

A microgrid consisting of hybrid renewable energy systems has been modeled by using a linear programming in [85] where the total annualized cost has been minimized by optimally selecting different system components and renewable resources for the smartgrid. The mixed integer linear program (MILP), where variables are either integer-value or real value, presented in [108] tackles the problem of optimal distributed generation by dividing the problem into two phases, namely site planning model and capacity planning model. The authors argued that their proposed model was computationally efficient with better optimal solution. In [59], a novel double-layer coordinated control approach for microgrid energy management is proposed. The proposed approach involves two layers: the schedule layer and the dispatch layer. The schedule layer is a MILP. The schedule layer provides an economic operation scheme based on forecasting data, while the dispatch layer provides power of controllable units based on real-time data. Energy management problem of a standalone MG integrated with RES has been formulated as MILP problem in [119] to minimize the operation cost and penalties related to unmet energy. In [144] a MILP is used to solve the multi-objective framework. The framework facilitates an efficient trade off between low operation cost and good energy services to the end users. The objective includes the cost of power exchange with the main grid, the startup and shutdown costs, the operating cost of distributed generators, the payment for demand response load, the penalty costs for involuntary load curtailment and renewable energy spillage. A generic mixed integer linear programming model has been proposed in [68] to minimize the operating cost of a residential microgrid. A mixed-integer linear programming model has been proposed in [13] for the optimal design and operation of an energy system consisting of combined heat and power. The proposed model has been used to formulate a multi objective function to minimize the capital and operational costs along with minimization of the carbon dioxide CO_2.

2.4.2.3/ Non linear programming

In [40], the power consumption model of transaction and communication based applications is presented into a Modified Simple Power Consumption model (MSPC) of a server. A non-linear program is used for the model. The proposed algorithm selects one of the servers for mixed types of applications so that the total power consumption of servers can be reduced. A mixed integer non-linear programming based computational framework has been proposed in [21] to evaluate the performance of a hybrid renewable energy system. A bi level programming approach is presented in [5] to analyze the competitive situations of hierarchical decision between an energy service provider (ESP) and a large central production unit. The ESP includes several microgrids, each one comprising of controllable loads and dispatched distributed generation units. The scheme is compared with the vertically integrated operation of the system, in which only one entity manages both the central production unit and the distributed resources of the microgrid. The automation level in a microgrid is a complex, non-linear and discrete optimization problem. The authors in [146] presented a procedure which is a scenario-based search method using local automation and remote control strategies in a combinatorial manner considering achievable benefits for each scenario. The authors simulated their proposed methodology on a sample microgrid. In [123], the problem is formulated as a non-linear mixed-integer program which minimizes capital and annual operational costs of renewable energy subject to a variety of system and unit constraints. With a broader use of distributed generation and storage, many buildings and campuses will become microgrids. Specific supply-side and demand-side aspects include on-site renewable generation, storage technologies, electric cars, dynamic pricing, and load management. A non-linear mixed integer optimization problem is used in [41] to solve these technical challenges.The authors formulated a long term scheduling optimization problem of a grid-tied renewable energy microgrid with hybrid energy storage as mixed quadratic program (MQP).

2.5/ Summary

This chapter helps to understand the main models used for each sources of the defined DATAZERO hybrid renewable energy sources. It is composed of photovoltaic panels, wind turbine, battery storage system and hydrogen storage system composed of fuel cell, electrolyzer and hydrogen tank in order to supply a data center of 1 MW peak consumption. The first section of this chapter defined the infrastructure of data center and their architecture. The maintained architecture of DATAZERO Project is A 2N data center. In the section 2.2, the mathematical models of the electrical sources were presented in details for two purposes; the model (A) is used for sizing and long term management and the model (B) is used for short control and management. Then, a state of art of sizing and management methodology were described in order to study the difference of between these methodology and to determine the best way for our infrastructure. The chapter 3 and chapter 4 presents respectively a hybrid novel method for the design and a linear programming for the management of the infrastructure.

Chapter 3

Sizing Strategy

In this chapter, the sizing strategy is explained and settled. The chapter starts by explaining the objectives and motivation needed to attend. Then, the principle of the sizing is detailed in order to understand the follo-up ideas. The third part of this chapter is dedicated to the methodology employed in order to design the hybrid renewable energy system chosen. To end up, the results obtained from the sizing tools are treated and explained using common applied performance metric.

Contents

3.1/ Objectives & Motivation

In the previous chapter, the modelings of the electrical sources used was presented for two different purposes. The first one is for sizing and management and the second one is for controlling the electrical sources. These models are used, in this chapter, in order to identify the production power produced by each sources for an horizon $\mathcal{H}$. This production coupled with an storage energy system (SES) must satisfy at best the data center demand ("Pload").

In fact, the objective of the sizing consist in finding the best size of the electrical sources used while guaranteeing the best answer to the power request (Pload) of the data center during a defined period. In fact, in case of good weather, the overproduction (when the renewable energy produced is bigger then the demand) is stored in the SES. In the opposite, in case of bad weather, the renewable power production can be not enough to supply the load. In these cases, storage capacity compensates the difference between the advertisement and the renewable production. Thus, the objective here is to use the renewable overproduction to fulfill the gap of underproduction during a defined period in order to ensure the green aspect and autonomy of the data center.

Consequently, a hybrid sizing method is presented in this chapter as we have chosen to place ourselves from an operational point of view where at each step of time we must wonder: how much energy must we load / unload? The aim is to have an electrical infrastructure that allows to respond to the data center power demand while using the renewable energy and keeping the same level of storage as at the beginning while respecting the defined performance criteria.

3.2/ Principle

In this section, the principle of sizing, on which the methodology is based, is explained. In order to understand the operating of the hybrid renewable energy system, the principle is divided into two different parts treating two type of infrastructure. The first one is composed of photovoltaic panels (PV) and hydrogen storage system. The second one is is constituted by the defined architecture of the project (photovoltaic panels, wind turbines, battery storage system, hydrogen storage system). **The mean average power during one hour is named power till the end of this thesis.**

3.2.1/ Photovoltaic and hydrogen infrastructure

To illustrate our general approach, we consider that the infrastructure is only composed of photovoltaic panels and hydrogen storage system. The PV are then the primary sources used to directly supply the data center demand. The hydrogen system is considered as a secondary sources used as a back up supply. In this case, we assume that, the idealized daily can be represented like the typical curve of the photovoltaic system. Moreover we assume that the production of energy by the primary sources during one year linearly increases when days get longer and linearly decreases when days get shorter (see Figure 3.1). In this figure n represents the number of days ($n = 365$) during a year, E and e are respectively the maximum and minimum production of energy in the year for each day. So 0 corresponds to December the 21th (the day with the shortest light) and $n/2$ to June the 21th (the day with the longest light).

During one year, the electrical production varies greatly over season. With making a zoom on the daily scale, it's obvious to differentiate the best production day (E) and the worst one (e) as showed in Figure 3.2 and where D is the power demand of the data-center. Considering our model, the overproduction of one day has to offset the underproduction of another day. And using both the wind turbine and the photovoltaic panel models introduced in chapter two, we can estimate their mean power production in kWh during these two days. This allows us to evaluate the primary source sizing. To make this sizing, we choose the worst and the best days (0 and $n/2$).

As mentioned before, the objective fixed is to fulfill the lack of energy of the worst day by the over-

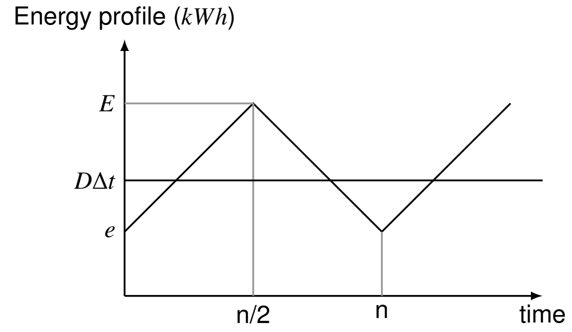


Figure 3.1: Primary energy produced by PV

production of the best day and to repeat this operation all over the year by taking the charging and discharging storage efficiency η into account. This problem is summarized by these equations (3.1):

$$\begin{aligned}(E - D\Delta t) \times \eta &= D\Delta t - e \\ \eta E + e &= D\Delta t(1 + \eta)\end{aligned} \tag{3.1}$$

where Δt is the time step ($\Delta t = 24\,\text{h}$). We assume the energy produced during the best day E and the energy produced during the worst day e could be written using a proportionality coefficient α: $e = \alpha E$.

$$E = \frac{D\Delta t(1+\eta)}{(\eta+\alpha)} \tag{3.2}$$

Then, the sizing of PV is made following the power demand that should be provided on the best day. Considering the value of E and the energy model for solar radiation, we can found the peak of power that the photovoltaic panels have captured to produce E at the end of the best day. Based on that, on can determine the surface of photovoltaic panels needed to produce E. However, it is necessary now to focus on the hydrogen storage system.

Taking storage efficiency into account, the former introduced model is not accurate anymore. The overproduction part needs to be bigger than the lack in the production part as it is multiplied by the storage efficiency. The real model then representing the system can be defined according to the scheme shown in Figure 3.3. The sizing of the hydrogen system consists in calculating the area of the triangle above the demand and is calculated by Equation (3.3) to get the energy needed to be provided in time of need:

$$E_{H2} = \frac{nD\Delta t(1-\alpha)}{4 \times (\eta+\alpha)} \tag{3.3}$$

Nevertheless, using this approach does not give enough information in order to design the real infrastructure composed of PV, WT as primary sources and Hydrogen and battery storage system as secondary sources. Moreover, one needs to take into account all the days over the year and not only be based on the best day in the idea of covering underproduction by overproduction. Thus, the main idea of covering underproduction by overproduction will remain in the following approaches and the the rest will not be taken into account.

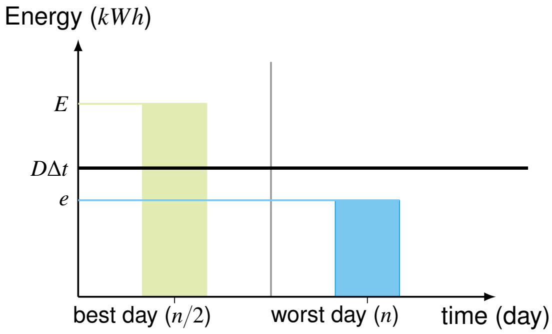


Figure 3.2: Primary energy produced by PV respectively for the best and the worst day (Integral of the power during both days)

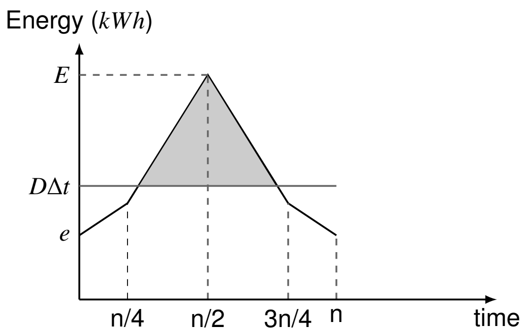


Figure 3.3: Required overproduction (gray part) for energy storage that takes efficiency ratio into account

3.2.2/ Main infrastructure

The main infrastructure of the hybrid renewable energy sources (HRES) is composed of photovoltaic panels (PV), wind turbines (WT), battery storage system (BSS) and hydrogen storage system (HSS). The aim of this infrastructure to supply the energy demand of the data center and to define the number of each source used. These electrical sources are divided into 2 different subsystems as explained in Figure 3.4:

- The primary sources: consist in providing the basic power to supply the data center and are composed of photovoltaic panels and wind turbines.
- The secondary sources: are the back up power to supply the data center in times of need and are composed of batteries and fuel cells.

As the data center should be autonomous in terms of energy consumption, the totality of the energy comes from primary sources. Moreover knowing that the primary sources operate as intermittent sources in time, we have to balance the lack of energy production (for example in the winter) by an over production (summer) during the year. To achieve this balance, the secondary sources (storage sources) are divided following the type of storage explained in the section above and operate as follows:

- Long-term storage: where days of overproduction will balance days of underproduction. The electrical resources used in this case is the hydrogen system;
- Short-term storage : where the hours of overproduction will balance the hours of under-

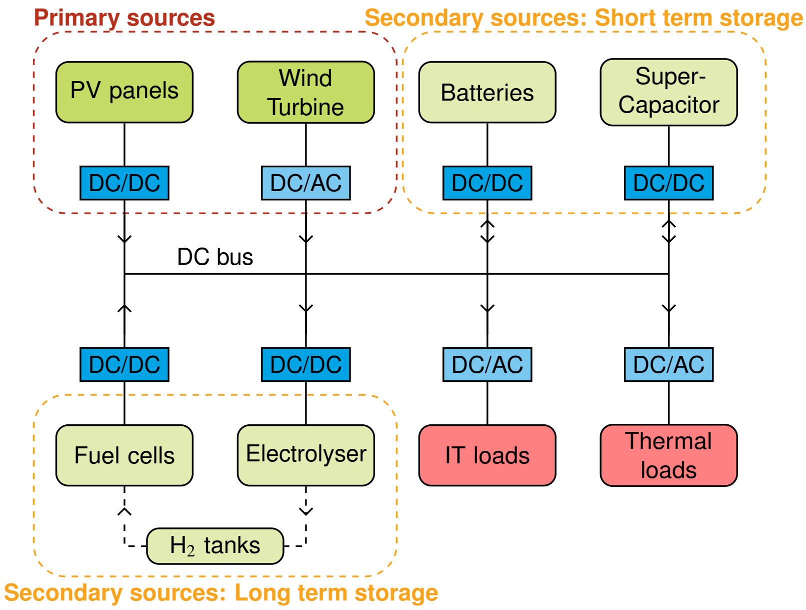


Figure 3.4: Subsystems of the HRES

production during the same day (fluctuations between day and night). It means that the production will be smoothed over the day. The electrical resources used in this case is the batteries.

In this section, the problem is taken backward.In fact, from a given infrastructure, the power load $Pprod$ is computed such as the overproduction is equal to underproduction. It is the available constant power that an infrastructure can provide all over the year. This is done in order to find a global calculation that determines a prediction model ensuring a cyclical use of the sources.

3.2.2.1/ Power load determination

In this part, the power load $Pprod$ is determined following a given primary production (PV + WT). Then the demand is computed such as the surface of the energy above it, multiplied by the storage efficiency η is equal to the surface below.

To make it possible, the renewable energy production is calculated based on the short model presented in the previous chapter. Then, all the primary production values x_i are sorted in a decreasing order. Then using binary search, we look for the exact value of k for which the overproduction is equal to the underproduction and calculate $Pprod$ following Equation 3.4. At the end, $Pprod$ is chosen according to the best combinations.

For instance, Figure 3.5 constitutes the power production by the renewable energy sources in each time slot X_1, X_2 till X_n, and $Pprod$ represents the exact value for which the surface (Energy) above $Pprod$ multiplied by the efficiency η is equal to the surface below $Pprod$. In other words, the overproduction is equal to the underproduction taking the storage efficiency η into account.

As a consequence, the general mathematical formula enabling the determination of the demand can be written as:

$$Pprod = \frac{\eta \sum_1^k x_i + \sum_{k+1}^n x_i}{n - k(1-\eta)} \tag{3.4}$$

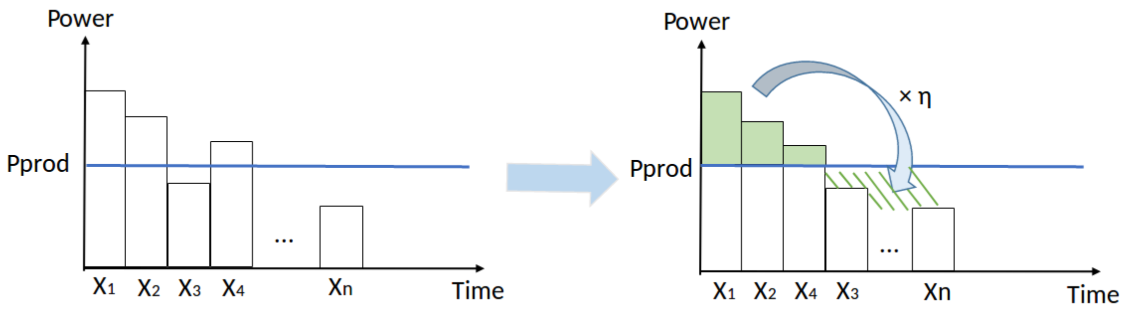


Figure 3.5: Compensation principle

where k is the number of time slots where the renewable energy production is bigger than the demand. Using Equation 3.4, one can determine for any renewable production, the exact maximal power value for which the overproduction is equal to the underproduction taking the storage efficiency into account. This maximal power value is the constant power that an infrastructure can provide all over the year Consequently, it reinforces the autonomous character of any infrastructure taken.

Nevertheless, the computing of $Pprod$ depends of the efficiency of the storage used. In this case, we are considering two different storage system. The approach is then explained in the two following paragraphs following the usage of the storage sources (short term or long term storage).

- **Short-term storage**

The short-term storage or in other words the daily offset consists in using the battery to cover the lack of production during hours within the same day. Then, the battery is used to smooth the production over the day. The calculation of the power needed to be stored or, in the other way, needed to be taken from the battery consists in computing $Pprod$ with the same formula in 3.4 for each day following the primary sources production taking the efficiency of the battery η_{batt} into account. Once computed, the result achieved consist in 365 values of $Pprod$, where each $Pprod$ represent the demand of one day of a year.

For instance, as explained in Figure 3.6 which represents only two days, let us suppose that X_1, X_2 are the production values for the first day, X_3,X_4 are production values of another day and that the efficiency of a charge discharge cycle of the battery $\eta_{batt} = 0.8$.

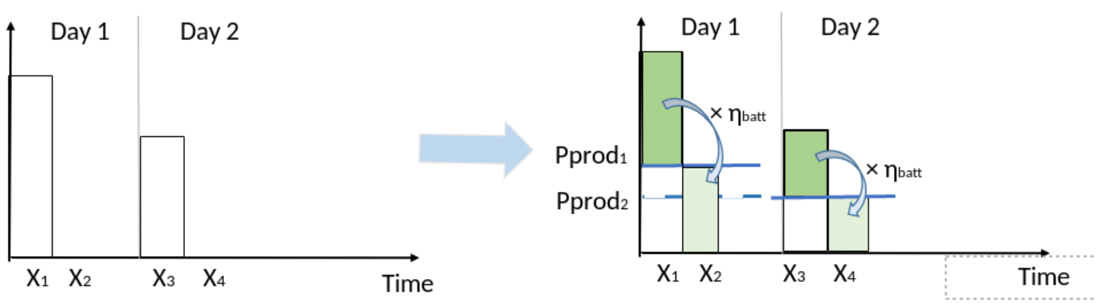


Figure 3.6: short term management in the same day

The production values of these time slot is as follows:

- $X_1 = 10$
- $X_2 = 0$

- $X_3 = 5$
- $X_4 = 0$

Then, following the same strategy explained above,

The energy production of the first day is computed as follows:

$$Pprod_1 = \frac{0.8 * 10}{2 - 1(1 - 0.8)} = 4.44$$

Second, we compute the energy production of the second day:

$$Pprod_2 = \frac{0.8 * 5}{2 - 1(1 - 0.8)} = 2.22$$

Using this strategy, the battery is smoothing the production during the same day. Then, this computing is made for every day over the horizon considered. The hydrogen is used to smooth the production over days. The used strategy is explained in the following section.

- **Long-term storage**

The long-term storage or seasonal offset consists in using the hydrogen system to compensate the lack of energy by using the fuel cell or also to store the overproduction of energy using the electrolyzer and store it in some hydrogen tanks to be used in time of need.

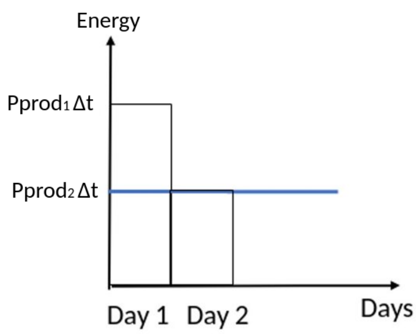


Figure 3.7: Double compensation process

The long term storage is calculated following the the 365 values of $Pprod$ obtained in the first part following the formula 3.4 using binary search. So, if we use the same example as before, we calculate again a new $Pprod'$ with Equation 3.4 using the efficiency of the hydrogen storage in this time as follows:

$$Pprod' = \frac{0.4 * (4.44 + 2.22)}{2 - (1 - 0.4)} = 1.90 \tag{3.5}$$

Where the efficiency of the hydrogen storage is the efficiencies of both fuel cell and electrolyzer $\eta_{H2} = \eta_{el} \times \eta_{fc} = 0.4$

With the long-term storage, a double compensation is made by smoothing the production, in the beginning, with battery and then with the hydrogen system. While computing the double compensation, we find out that it's **the lower bound** quoted Lb that we can obtain for the power load $Pprod$.

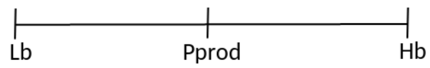


Figure 3.8: Dichotomy calculation for the power load of an infrastructure

Based on this strategy, the constant power load, that an infrastructure is able to deliver, is between a lower bound and a higher bound. Moreover, if the renewable power produced during a day is bigger then the power demand, it is better to store the overproduction directly in hydrogen and not into the batteries to avoid losses of energy. Therefore, in the following part, the computing of the **higher bound** is explained.

The higher bound quoted Hb consists in computing the average power for each day. As a result, we will get 365 values of daily average. With these values, we calculate again the yearly average. This bound represents the upper bound because with computing the mean of the daily average, we do not take into account neither the efficiency of the battery nor the one of the hydrogen system.

For instance, the higher bound is computed as follows:

$$M1 = \frac{10+0}{2} = 5 \qquad M2 = \frac{5+0}{2} = 2.5 \tag{3.6}$$

$$Hb = \frac{5+2.5}{2} = 3.75 \tag{3.7}$$

- **Conclusion:**

To conclude, the aim is to find more precisely the value of the load that a hybrid infrastructure is able to produce over the year. Moreover, this value is between the two limits: the higher and lower bound respectively Hb and Lb. Using these 2 values the power load $Pprod$ is computed using binary search between these two bounds and the storage process (which is explained in the following sections). The right value of demand is reached when the level of hydrogen, after the defined horizon, attains the same initial level at the beginning.

Using this strategy, the renewable energy produced from any infrastructure including hydrogen and battery storage system can define an exact constant power load for which the HRES can be completely reliable over the horizon considered.

3.3/ HRES sizing methodology

In this section, sizing methodology is presented. In the beginning, in order to better understand the operating of a hybrid infrastructure, we proceeded to a power load determination using a calibration of the renewable energy production. In the section 3.3, the data center load is defined. Then, our methodology consist in finding the best sizing for the client demand.

3.3.1/ Sizing of the storage system:

The strategy that aims to design the storage system rely on computing time slots of overproduction and underproduction for each day during the time horizon. Thus, from the renewable energy production, the sizing of the batteries and the hydrogen system can be started. As a reminder, the hydrogen system composed of fuel cells and electrolyzers is used to balance day of overproduction and days of underproduction (seasonal offset). However, the batteries are used during the same day to balance the hours of overproduction and underproduction (fluctuations between day and night).

This sizing strategy is explained in the two diagrams 3.10, 3.12. It starts by initializing the power values respectively of the battery P_{batt} and the hydrogen P_{H_2}, the Energy that is Needed to be Stored (ENS) in the batteries to handle the daily smoothing and the energy already stored in the batteries $Estored_{batt}$. Variables i and d represent respectively hours and days.

Second, to get the exact quantity needed to be stored in the storage system and the quantity needed to be taken from it in time of need, we compute the missing energy (underproduction Up) and the overabundant energy (overproduction Op) during a day with a time step of one hour as defined in equations (3.8), (3.9):

$$Overproduction : Op = \sum_{k=0}^{24}(x_k - D_k) \times \Delta t \quad for\ x_k > D_k \tag{3.8}$$

$$Underproduction : Up = \sum_{k=0}^{24}(D_k - x_k) \times \Delta t \quad for\ x_k < D_k \tag{3.9}$$

where x_k is the renewable production power over the hour k and D_k is power demand in the same hour. Thus, based on these two energy quantities Op and Up, a day can be defined as an overproduction or a underproduction day taking the storage efficiency into account as explained in the table 3.1:

Table 3.1: days management

Overproduction day	***Underproduction day***
$Op \times \eta_{batt}^2 > Up$	$Op \times \eta_{batt}^2 > Up$

Where the definition of the battery efficiency has changed: it is the charging or discharching efficiency. This explains the square η_{batt}^2 in the table 3.1. First let's start by explaining the sizing strategy of the storage system in an overproduction day.

- **<u>Overproduction day:</u>**

The sizing of the battery is a pessimistic sizing. In days of overproduction as showed in the figure 3.9, the energy needed to be stored in the batteries ENS is equivalent to the underproduction hours taking into account the battery efficiency. By this way, the batteries smooth the day. Then, the remaining overproduction energy is stored in the hydrogen system. To make it possible, hours of overproduction should be differentiated of hours of underproduction. To summarize, in an overproduction day, the steps are (from step 5 in 3.10):

- **Step 5:** As the batteries should assure the daily smoothing, the energy that should be stored from overproduction hours is equal to the sum on the underproduction hours. This energy is defined as:

$$ENS = \frac{Up}{\eta_{batt}^2}$$

- **Step 6:** Check if it's an underproduction or overproduction hour
- **Step 7:** In case it's an overproduction hour, the energy management of the storage system is defined as follows:
  - **Step 8-1:** Verify if the energy that is needed to store ENS in the batteries is smaller than the renewable energy produced in the same hour $Op[i]$

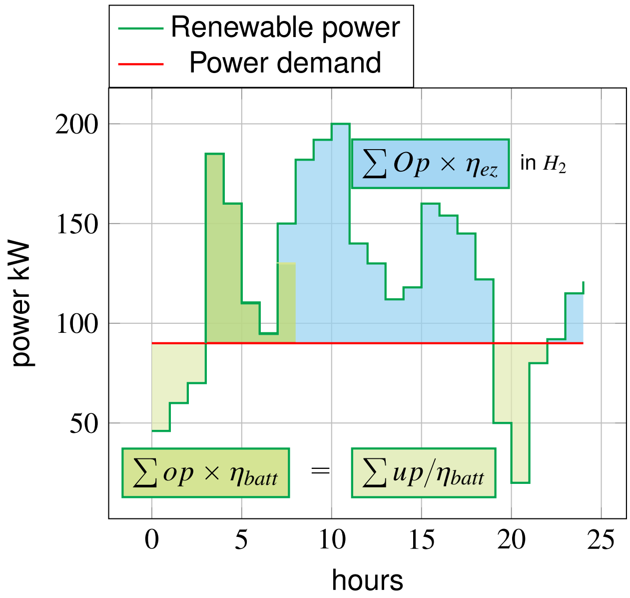


Figure 3.9: Overproduction day

- **Step 9-1:** If the latter condition is true, the battery charge all the energy that is needed to store ENS
- **Step 10-1:** The hydrogen charge the remaining energy of the renewable energy produced taking the electrolyzer's efficiency η_{elec} into account
- **Step 11-1 : Step 13,1:** Initialize the variable of the energy needed to store and increment the counter of hours to evaluate the next hour of the day

- **Step 8-2:** Verify if the energy that is needed to store ENS in the batteries is bigger than the renewable energy produced in the same hour $Op[i]$
- **Step 9-2:** If the latter condition is true, the battery charge all the energy that is produced by renewable energy in the same hour $Op[i]$
- **Step 10-2:** As there is no remaining energy, the hydrogen does not charge
- **Step 11-2:** Decrease the energy needed to store by what is already charged in the batteries
- **Step 12:** Test the hour counter and increment it if i is less than 24 and reconsider for the next hour.

• **Step 7-3:** In case of underproduction hour, the energy management of the storage system is defined as follows:

  - **Step 8-3: 9-3:** Discharge the battery of the energy of underproduction taking the efficiency into account
  - **Step 10-3:** As the hydrogen is responsible of the seasonal offset, the fuel cell does not start here.
  - **Step 11:** Test the hour counter and increment it if i is less than 24 and reconsider for the next hour.

• **Underproduction day:**

In an underproduction day as showed in the figure 3.11, only the energy equivalent to the hours of overproduction is stored in the batteries to supply the underproduction hours of the same day. To provide the remaining quantity of underproduction, fuel cells start working to handle the data center demand. This operations are summarized in the diagram 3.12 and explained as follows:

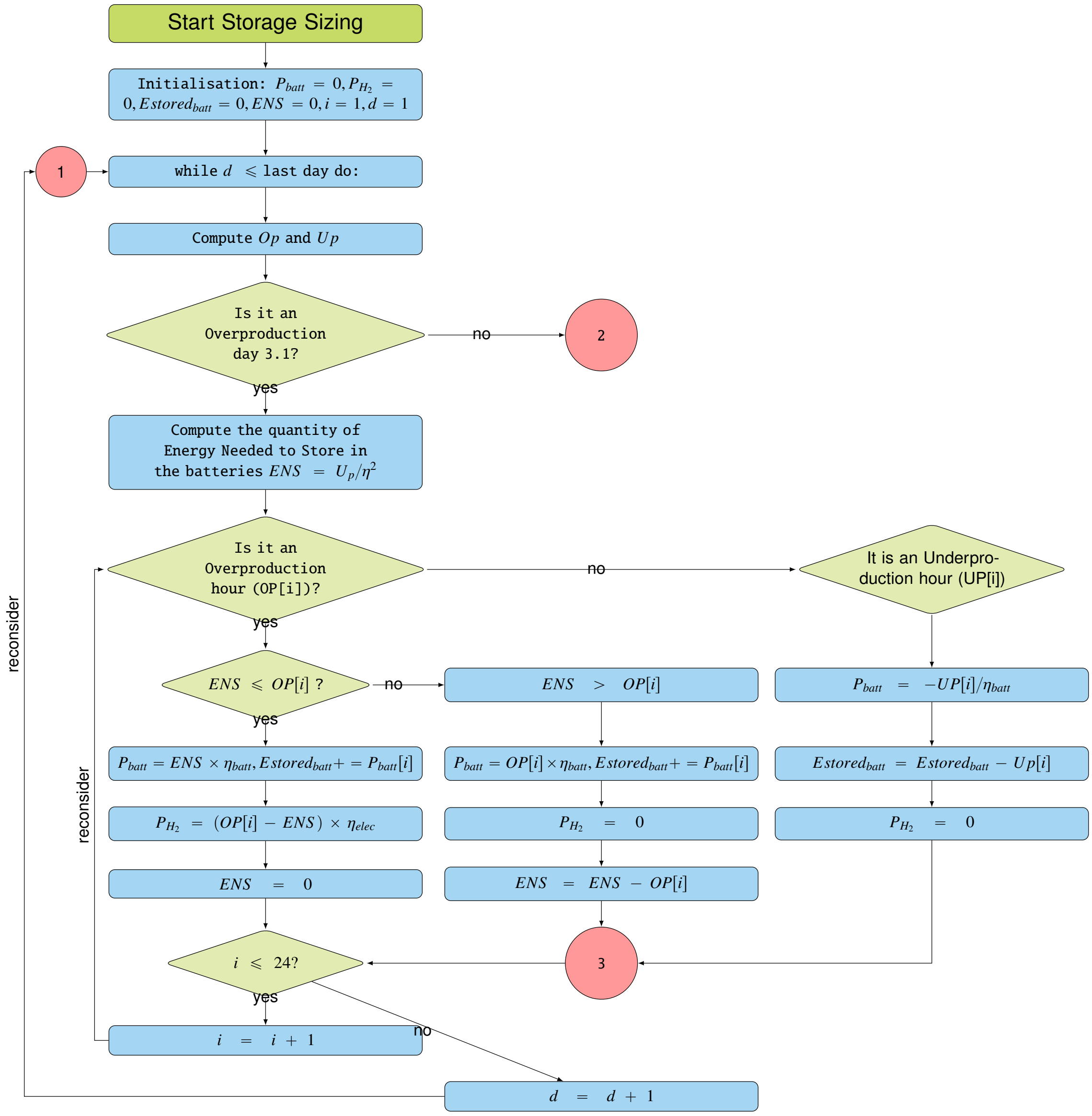


Figure 3.10: Overproduction day sizing

- **Step 2:** It's an underproduction day
- **Step 3:** As the batteries should assure the daily smoothing, the energy that should be discharged is equivalent to the overproduction hours. The latter is defined as :

$$END = Op \times \eta_{batt}^2$$

- **Step 4:** In case it's an underproduction hour, the energy management of the storage system is defined as follows:
  - **Step 5-1:** Verify if the energy that is needed to discharge END in the batteries is smaller than the renewable energy produced in the same hour $Up[i]$
  - **Step 6-1:** If the latter condition is true, the battery discharge all the energy that is needed to discharge END and could not be able to meet the data center demand.
  - **Step 7-1:** The hydrogen discharge the remaining energy to supply the data center demand taking the fuel cell's efficiency η_{fc} into account

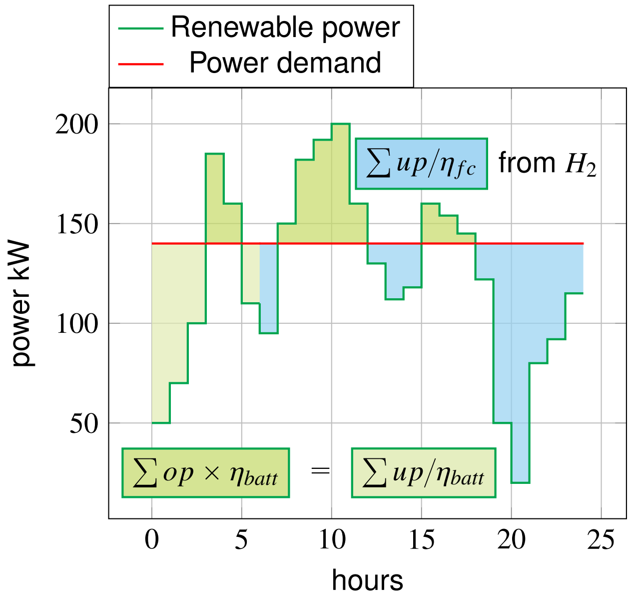


Figure 3.11: Example of a underproduction day

– **Step 8-1:** Initialize the variable of the energy needed to discharge and go to the point 3 to pass to the next hour/day.

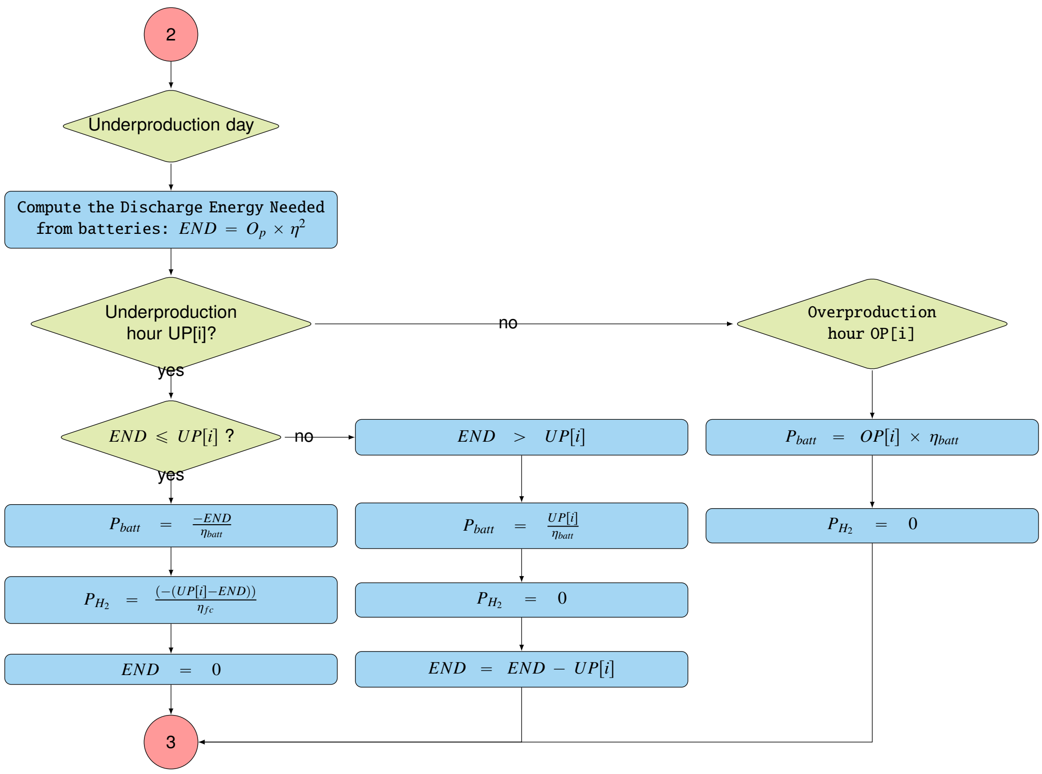


Figure 3.12: Underproduction day sizing

– **Step 5-2:** Verify if the energy that is needed to discharge END from the batteries is bigger than the underproduction $Up[i]$ of the same hour.

  - **Step 6-2:** If the latter condition is true, the battery discharge all the energy that is needed to supply the demand of the data center $Up[i]$ in the same hour.
  - **Step 7-2:** As there is no a remaining energy, fuel cells does not work.
  - **Step 8-2:** Decrease the energy needed to discharge by what is already discharged from the batteries and go to the point 3 to pass to the next hour/day.

- **Step 4-3:** In case of overproduction hour, the energy management of the storage system is defined as follows:
  - **Step 5-3:** charge the battery of the energy of overproduction taking their efficiency into account
  - **Step 6-3:** As the hydrogen is responsible of the seasonal offset, the fuel cell does not start here. Then go to the point 3 to pass to the next hour/day.

- **Synthesis**

In order to balance the lack of energy production by an over production during the year, taking the intermittent nature of the primary into account, we divided the storage as follows:

- Long-term storage: where the day of overproduction will balance the days of underproduction. The electrical resources used in this case is the hydrogen system
- Short-term storage : where the hours of overproduction will balance the hours of under-production during the same day (fluctuations between day and night). It means that the production will be smoothed over the day. The electrical resources used in this case are the batteries.

Using the strategy explained in the subsection above, the battery will always come back to the initial capacity after 24 hours of the starting time. Moreover, at the end of the whole period, (fixed in simulation for a year) the hydrogen also comes back to its initial state at the beginning of the period. Once the strategy of sizing the back up sources is fixed, one needs to design the primary sources which is the aim of the following section.

3.3.2/ Sizing of the primary source

Taking the demand of the data center into account, we can choose different way of sizing. To sweep up all the cases, in the beginning, the wind turbines are maximized and the difference is filled by photovoltaic panels. Then, the same calculation is made with one wind turbine less and more photovoltaics. As a result, we obtain different configuration of the hybrid renewable energy system.

First of all, let $\mathcal{H} = K\Delta t$ a time window discretized into K identical periods of duration Δt unit of time. Depending on the case, let k defined as both the kth time step at instant $k\Delta t$ from the beginning of $\mathcal{H}$ and the index of the kth period Δt of $\mathcal{H}$ defined between the two time step k and $k+1$ ($k \in [\![0, K]\!]$)
.

3.3.2.1/ Sizing of the wind turbine

First, we compute the energy produced by the wind turbine during one year with a time step of one hour using the equation(3.10):

$$E_{wt} = \sum_{k=0}^{K} P_w(k) \times \Delta t \tag{3.10}$$

Then, we compute the maximum number of power turbine that could be used in the installation following the equation (3.11) where E_D is the energy demand of the data center during an horizon $\mathcal{H}$:

$$Nmax_{wt} \geqslant \left\lceil \frac{E_D}{E_{wt}} \right\rceil \tag{3.11}$$

3.3.2.2/ Sizing of the photovoltaic panels

First, we compute the energy produced by one m^2 of photovoltaic panel during an horizon $\mathcal{H}$ using the equation (3.12):.

$$E_{pv} = \sum_{k=0}^{K} Ppv(k) \times \Delta t \tag{3.12}$$

Second, as the maximal number of wind turbine is defined, we use binary search to define the exact surface A_{pv} of photovoltaic panels that is needed for the infrastructure. In fact, the exact surface is the one for which the computed level of hydrogen $LOH(K)$ at the end of the horizon $\mathcal{H}$ comes back to the same initial values $LOH(0)$. It is computed such as:

- if $LOH(K) > LOH(0)$, the surface should decrease;
- if $LOH(K) < LOH(0)$, the surface should increase;
- if $LOH(K) = LOH(0)$, the right value for the surface of PV is found.

This design can be defined in the following algorithm 1:

Algorithm 1: Computation of the surface Apv of solar panels

Require: $Nmax_{wt}$, the charge $D(0..K-1)$ along the horizon $\mathcal{H} = K \times \Delta t$
  for $nbWT = 0 : 1 : Nmax_{wt}$ **do**
    $Apvmax \leftarrow MAX\_SURFACE$ {Maximal surface}
    $Apvmin \leftarrow 0$ {Minimal surface}
    $Pw(k) \leftarrow P_w(k) \times nbWT \forall k \in [\![0, K-1]\!]$
    $LOH(0) \leftarrow LOHinit$
    $LOH(k) \leftarrow 0 \forall k \in [\![1, K]\!]$
    while $Apvmax - Apvmin \geqslant 1 \quad and \quad LOH(H) \neq LOH(0)$ **do**
      $Apv(nbWT) \leftarrow (ApvMax + ApvMin)/2$
      $Ppv(k) \leftarrow I(k) \times Apv \times \eta_{pv} \forall k \in [\![0, H-1]\!]$
      $Pre(k) \leftarrow Ppv(k) + Pw(k) \forall k \in [\![0, H-1]\!]$
      $LOH \leftarrow \texttt{proceedSizingStorage}(D, Pre)$
      if $LOH(H) \leqslant LOH(0)$ **then**
        $Apvmin \leftarrow Apv(nbWT)$
      else
        $ApvMax \leftarrow Apv(nbWT)$
      end if
    end while
  end for
  return $Apv(0..Nmax_{wt})$ the $Nmax_{wt} + 1$ configurations

The “proceedSizingStorage” in the algorithm 1 is the computation of sizing storage while taking into account the daily and seasonal compensation as explained in the precedent subsections over the year. For instance, if the chosen horizon is equal to one year, then, $\mathcal{H}$ is equal to 8760 and $Deltat = 1$ h. Thus, the “proceedSizingStorage” algorithm level turns on all $\mathcal{H}$, but with an internal treatment 24h. In fact, it is composed of an internal loop turning on 24h (24 iterations) and an external loop running on 365d (365 iterations)

Then, we repeat the same sizing strategy with one power turbine less and the equivalent surface of PV calculated till reaching the all PV configuration.

As a result, we obtain different configurations of the hybrid renewable energy system. Nevertheless, in front of many optimal solutions, one needs to settle some metrics in order to choose the best one.

3.4/ Applied performance criteria

In order to analyze the results, several metrics have been settled in order to quantify the obtained solutions.

3.4.1/ Annual system cost

The Annualized Cost of the System (ACS) includes the capital cost (C_{cap}), the replacement cost in case of system breakdowns (C_{rep}) and operation and maintenance cost to ensure the lifetime of the system($C_{O\&M}$) defined by:

$$ACS = C_{cap} + C_{rep} + C_{O\&M} \tag{3.13}$$

- The capital cost (C_{cap})

$$\begin{aligned} C_{cap} = {} & Ccap_{PV} \times Pmax_{PV} + Ccap_{WT} \times P_r + \\ & Ccap_{batt} \times C_{batt} + Ccap_{fc} \times Pmax_{fc} + \\ & Ccap_{el} \times Pmax_{el} + Ccap_{H_2} \times C_{H_2} + Ccap_{inv} \end{aligned}$$

where: $Ccap_{PV}$, $Ccap_{WT}$, $Ccap_{batt}$, $Ccap_{fc}$, $Ccap_{el}$, $Ccap_{H_2}$ and $Ccap_{inv}$ the capital cost of each sources as explained in the 3.3. Ppv_{max} , P_r, Pfc_{max}, Pel_{max} are the maximal power attained by the photovoltaic panel, wind turbine, fuel cell and the electrolyzer. Finally, C_{H_2} and C_{batt} are the capacity of the battery and the hydrogen tank. We define $\mathcal{S} = \{PV, WT, batt, fc, el, H_2\}$ is the set of the electrical sources used with $s \in \mathcal{S}$.

- The replacement cost (C_{rep})

$$C_{rep} = \sum_{s \in \mathcal{S}} Crep_s \times Pmax_s \times \frac{T_{sys}}{T_s} \tag{3.14}$$

where $Crep_s$ is the replacement cost of each one of the resources used s with a defined lifetime T_s, $Pmax_s$ is the maximal power delivered by the sources. Additionally, we defined the operating time of the hybrid system (T_{sys}).

- The operation and maintenance cost ($C_{O\&M}$)

$$C_{O\&M} = \sum_{s \in \mathcal{S}} Cmain_s \times Pmax_s \times T_s \tag{3.15}$$

Where $Cmain_s$ is the operation & maintenance cost of each resource s used with a defined lifetime T_s, $Pmax_s$ is the maximal power delivered by the sources.

3.4.1.1/ Reliability

Due to the intermittent nature of the power produced by PV and WT, analyzing the reliability of the system is quite important. Actually, an electrical system is considered reliable when it is able to supply the required power. There are many methods for carrying out the reliability analysis: The Loss of Power Supply Probability (LPSP) defined in the following equation is one of the most used ones:

$$LPSP = \frac{\sum_{k=0}^{K-1} \mathbb{1}_{Pavai_k < D_k}(k)}{K} \tag{3.16}$$

where $Pavai_k$ and D_k are respectively is the available power and the power demand in the k^{th} time step Δt. $\mathbb{1}_{Pavai_k < D_k} = 1$ if the condition $Pavai_k < D_k$ is true, 0 otherwise.

$$Pavai_k = Pre_k + Pbatt_k + Ph2_k \tag{3.17}$$

with Pre is the renewable power produced, $Pbatt$ and $Ph2$ are respectively the charging or discharging power of the battery and the hydrogen system $\forall k \in [\![0, K-1]\!]$.

A loss of power suplly equal to zero means that the system is absolutely reliable and a LPSP of one means that the load is not satisfied at all by the generated power.

3.4.1.2/ level of autonomy

The level of autonomy (LA) is defined as the percentage of hours where the demand is satisfied by the renewable energy solely. It is computed as defined in the equation 3.18.

$$LA = \frac{\sum_{k=0}^{K-1} \mathbb{1}_{Pre_k \geqslant D_k}(k)}{K} \times 100 \tag{3.18}$$

Where $\mathbb{1}_{Pre_k \geqslant D_k} = 1$ if the condition $Pre_k \geqslant D_k$ is true, 0 otherwise.

3.5/ Results and Analysis

In this section, the input data to the program (such as the meteorological data, the efficiencies values, etc.) are given. The obtained results are explained in a second time in two paragraphs. The first one treats the power load determination approach. Its aim is the find the exact power load that an infrastructure can provide all over the year. The second paragraph is dedicated the results of the sizing methodology for a given load of the data center.

3.5.1/ Input data for simulations

- **Meteorological and IT data**

To simulate the primary sources production (photovoltaic panels and wind turbines), one needs to download meteorological data such as the irradiance and wind speed data. These data can be obtained from various database online. On our case, the irradiance data are downloaded from the National Solar Radiation Database (NSRDB) [116], and the wind speed data are downloaded from the wind prospector from the national Renewable Energy Laboratory (NREL) [43]. These data can be taken every hour during a day starting from 2004 till 2012. The chosen localization is Los Angeles with the details of Table 3.2.

Table 3.2: Meteorological data localization

Latitude	***Longitude***	***Time Zone***	***Elevation***
34.57	-118.02	-8	807

On the other side, for the results methodology, the load used to simulate the data center is the traces of access logs from the World Cup web site of 1998 [23]. These meteorological data and IT traces are used as main input of the sizing strategy programmed in Python 3. In the subsection 3.5.2, the traces are not taken into account as it aims to determine a power load for which the HRES is completely autonomous.

- **Electrical data**

The electrical specification of the hybrid renewable energy system are defined in Table 3.3. All these specifications are used as inputs for the program. As a result, one can obtain different configurations of sizing of the hybrid energy system composed of photovoltaic panels, wind turbines, batteries and hydrogen system for each year.

3.5.2/ Power load determination results

Several years of meteorological data (from 2004 till 2012) are taken into account in order to verify the robustness of the approach.

As a reminder, in this part, the renewable energy produced by both wind turbines and photovoltaic panels is computed during the whole year (8760 hours). These values are then sorted out in a decreasing order to determine the power value for which the overproduction is equal to the underproduction taking the storage efficiency into account. Then the management of storage energy system is played where the batteries ensure the short term storage (daily smoothing) and the hydrogen system ensure the long term storage (seasonal offset).

- Battery storage system: the capacity of the battery must be in the same level as the initial capacity as it only needed to the daily compensation. It means that the energy charged in the batteries are discharged in the hour of need and the battery comes back to the same level as the beginning of the day.
- Hydrogen storage system : the level of hydrogen (LOH) at the end of the horizon defined, 1 year in our simulation example, should be equal to the initial LOH. It means that:

$$LOH_K = LOH_0 = LOHinit \tag{3.19}$$

Once all the scenario is played, the power load value, that the system is capable to ensure all over the year, is computed and displayed.

For instance, the meteorological data of 2005 for Los Angeles are used with an hourly timescale to verify the correct operational state of the equipment. Moreover, the 2N data center infrastructure must be able to deliver 1MW in time of peaks. Given the redundant nature of the infrastructure, each branch (N) of the data center should be capable of delivering peak power equal to 500 kW. Thus, the infrastructure was sized and calibrated such as it is capable to deliver a constant power load each hour equal to 500 kW.

Table 3.3: Nomenclature

Variable	Description
Specifications: Solar panel	
Rated module efficiency η_{pv}	$15\,\%$
Lifetime	25
Efficiency	$15\,\%$
Capital cost $Ccap_{pv}$	2000 per kW
O&M cost $Cmain_{pv}$	30 per kW per year
Replacement cost $Crep_{pv}$	1200 per kW
Specifications: Wind turbines	
Rated power P_r	500 kW
Cut-in speed v_{ci}	$4m/s$
Rated output speed v_r	$14m/s$
Cut-out speed v_{co}	$25m/s$
Lifetime	20
Capital cost $Ccap_{wt}$	3000 per kW
O&M cost $Cmain_{wt}$	40 per kW per year
Replacement cost $Crep_{wt}$	1500 per kW
Specifications: Batteries	
Battery charge efficiency η_{ch}	90 [%]
Battery discharge efficiency η_{dch}	90 [%]
Lifetime	20
Capital cost $Ccap_{batt}$	600 per kW
O&M cost $Cmain_{batt}$	5 per kW per year
Replacement cost $Crep_{batt}$	420 per kW
Specifications: Hydrogen tank	
Lifetime	30
Capital cost $Ccap_{H_2}$	600 per kW
O&M cost $Cmain_{H_2}$	5 per kW per year
Replacement cost $Crep_{H_2}$	420 per kW
Specifications: Electrolyzer	
Electrolyzer efficiency η_{ez}	60 [%]
Lifetime	40000 hours
Capital cost $Ccap_{el}$	2700 per kW
O&M cost $Cmain_{el}$	30 per kW per year
Replacement cost $Crep_{el}$	1600 per kW
Specifications: Fuel Cell	
Fuel Cell efficiency η_{fc}	60 [%]
Lifetime	40000 hours
Capital cost $Ccap_{fc}$	4000 per kW
O&M cost $Cmain_{fc}$	30 per kW per year
Replacement cost $Crep_{fc}$	2500 per kW
Specifications: other	
Inverter efficiency η_{inv}	[%]
Lifetime	15 years
Capital cost $Ccap_{inv}$	600 per kW
O&M cost $Cmain_{inv}$	0 per kW per year
Replacement cost $Crep_{inv}$	600 per kW
System lifetime $Tsys$	30 years

As a result of the simulation, different sizing configurations capable of delivering variable power load are summarized in the Table 3.4.

The different sizing configuration of the year 2005 are detailed as follows:

- Full PV: this configuration is composed by only photovoltaic panel as renewable sources to deliver the power load. In this configuration the storage is bigger as it must ensure the demand during nights where there is no solar.

Table 3.4: Results of the sizing tool for the year 2005

Nwt	**Npv** (m^2)	**Battery capacity** (kWh)	**Hydrogen size** (kg)	**Demand** (kWh)
0	18269	8027.56	15511.10	500.02
1	11720	6927.48	10301.83	499.98
2	7011	5546.82	8026.92	499.99
3	3050	5649.51	9581.05	499.98
4	0	8603.93	12241.93	517.32

- Hybrid configuration: these configurations are composed by an increasing number of wind turbine with respectively a decreasing number of photovoltaic panels.
- Full WT: this configuration is composed by only wind turbine as renewable sources to supply the data center demand. One can see that the demand delivered by this infrastructure is bigger than the peak demand. This is due to the rated power of the turbine.

All these configuration delivers as expected 500 kW each hour. However, the Full WT configuration delivers more than expected as the rated power of the wind turbine is the constant and same for all configurations. The configurations obtained for the different years for 2004 till 2012 are in the appendix A. This table summarizes the sizing infrastructure for each year of the data base able to deliver as a constant power load $Pprod$ equal to 500 kW. The results are different since each of these years has different sunshine and wind profiles.

In the same time, common sens rules detailed in the sections before can be illustrated in Figure 3.13, where the illustrated sizing configuration is 2 wind turbine and 8762 PV panels. In this figure, one can see that the renewable energy is maximized during the day and decrease during nights. This behavior is completely normal as this sizing prioritize the PV panels. The demand (Load in the first sub-figure) obtained by binary search from this renewable profile is equal to $Pprod = 500.02$ kW equal to the data center demand $Pprod = 500$ kW. One can see also that the level of hydrogen at the end of the horizon $\mathcal{H}$ (equal to 8760 hours in this case) comes back to the initial level at $k = 0$. This behavior is conform with the strategy of the long term storage where the hydrogen is supposed to maintain the seasonal variation. The hydrogen power in red in the sub-figure represents both power of the fuel cell and the electrolyzer. When the latter is positive, it represents the power of the electrolyzer and when it's negative it represents the power of the fuel cell.

However, in order to show the proper use of the battery, Figure 3.14 shows a zoom of the Figure 3.13 during the first 10 days (240 hours). In fact, the capacity of the battery comes back to its initial level after each 24hours to ensure the day and night alternation.

This sizing method can calibrate thus any hybrid renewable energy system in order to make it 100 % autonomous while delivering a certain load for each hour.

3.5.3/ HRES infrastructure results

In this part, the same years of meteorological data from 2004 till 2012 for Los Angeles are used as an illustrative case and with a simulated load. It means that the obtained results varies with another load or meteorological data.

For instance, one can see the different configurations obtained for the year 2004 in Table 3.5.

Figure 3.15 shows the different configuration behavior related to a specific sizing. In the same time, the rules detailed in the sections before are maintained and are illustrated in Figure 3.15. In this figure, the behavior of the resources used in each configuration during 8760 hours of the year 2004 of simulation are shown. Figure 3.15a is composed of a full PV panel system. First, one can see the capacity of the battery which comes back to the initial state for each 24 hours. Thus, the daily smooth (day and night alternation) is maintained and ensured. Second, the variation of the level of hydrogen in the tank can be seen. As the simulation starts in January (winter), the level of hydrogen is decreasing because there is not enough sun to maintain the data center demand.

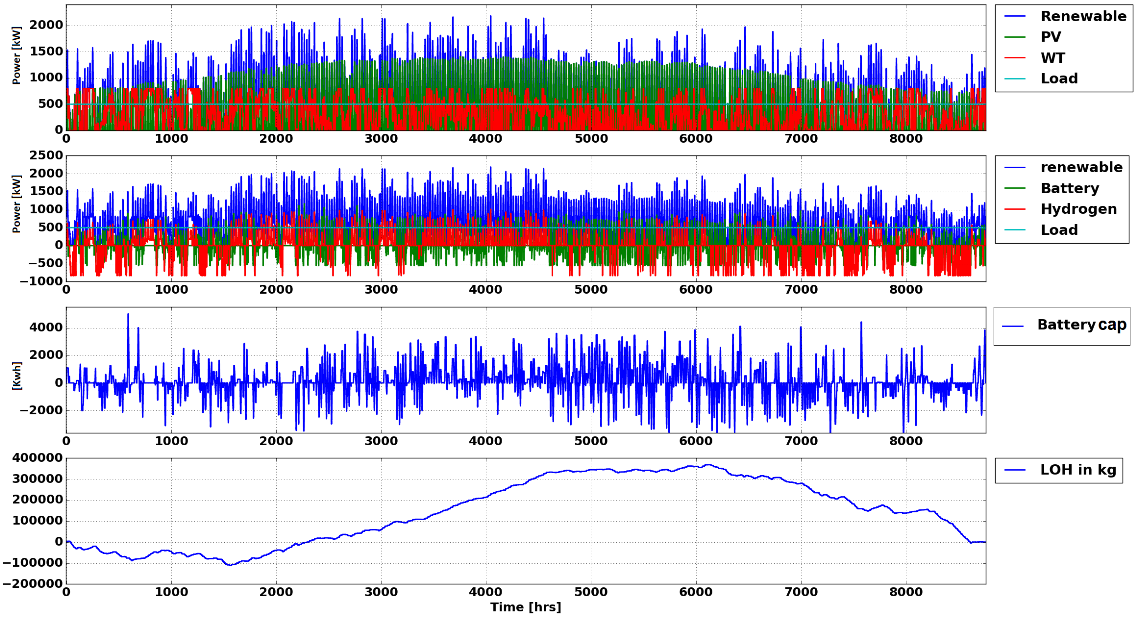


Figure 3.13: Result of the sizing management of the year 2005 with 2 wind turbine and 8762 pv panels

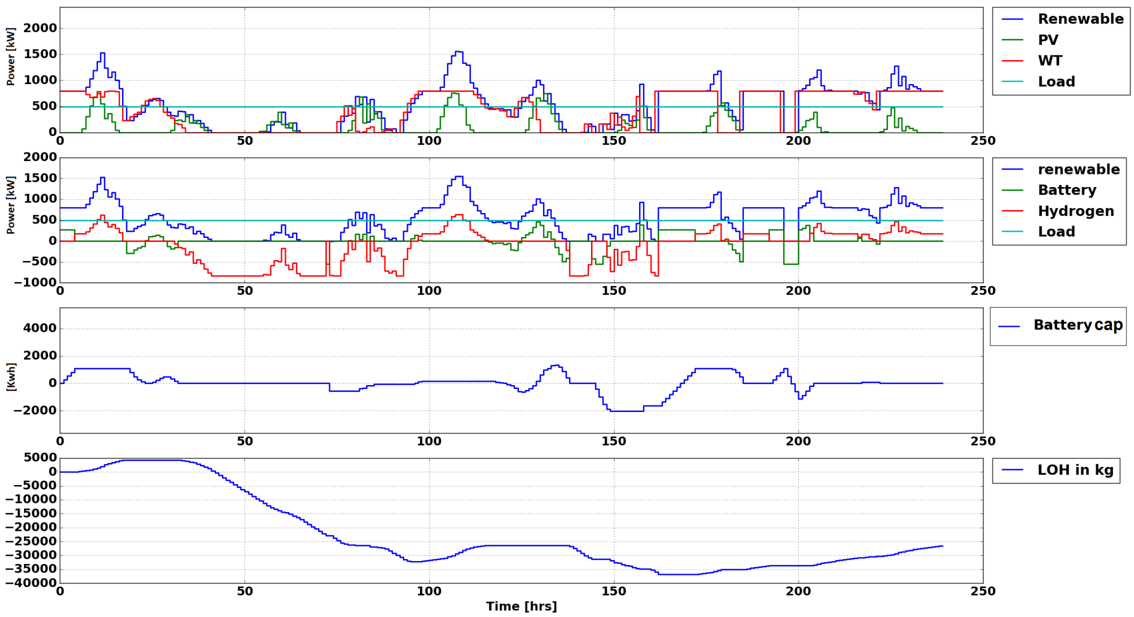


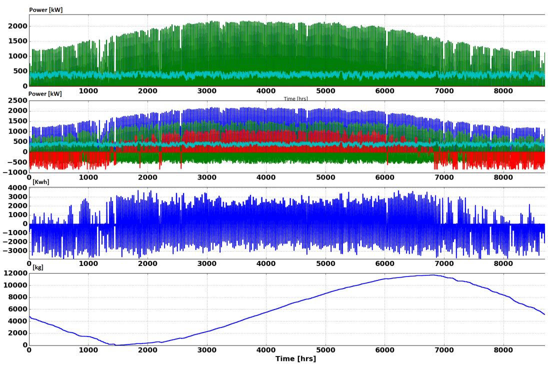
Figure 3.14: Result of the sizing management of the year 2005 with 2 wind turbine and 8762 pv panels (Zoom)

Then, it starts charging as it is the summer period until reaching the initial level of charge in the beginning of the period fixed (1 year).

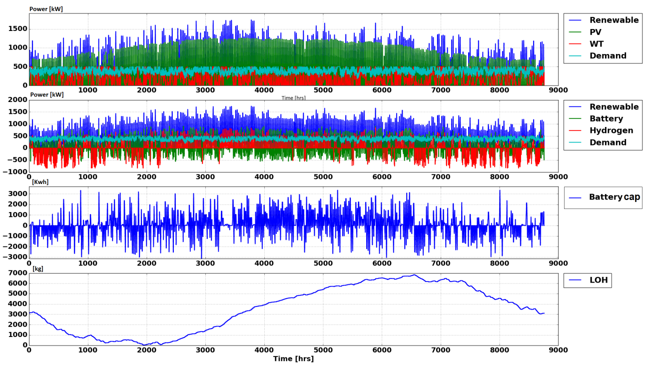
In Figures 3.15b and 3.15c, 3.15d, one can see the energy produced by wind turbine and PV panels. The common sens rules are maintained where also here the battery assures the daily smoothing and the hydrogen assures the seasonal offset. Nevertheless, some differences can be remarked compared to the previous figure:

Table 3.5: Results of the sizing tool for the year 2004

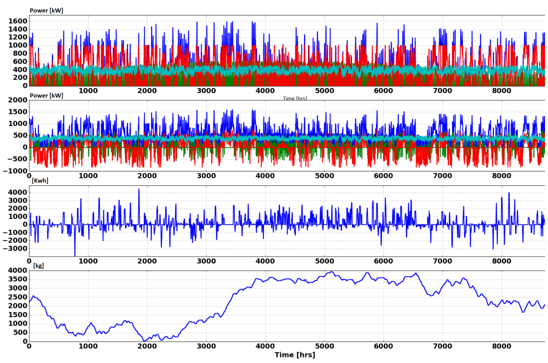
Config	*Nwt*	*Npv*	*Battery capacity (kWh)*	*electrolyzer(kW)*	*Fuel cell(kW)*	H_2 *size (kg)*
A (3.15a)	0	13587	7035.92	1115.16	832.65	11720.40
B (3.15b)	1	7883	4935.8	840.76	832.65	6845.29
C (3.15c)	2	3791	4466.65	779.85	832.65	3935.54
D (3.15d)	3	446	6110.37	798.77	832.65	4697.23
E (3.15d)	4	0	6435.36	1075.785	832.65	17920.35



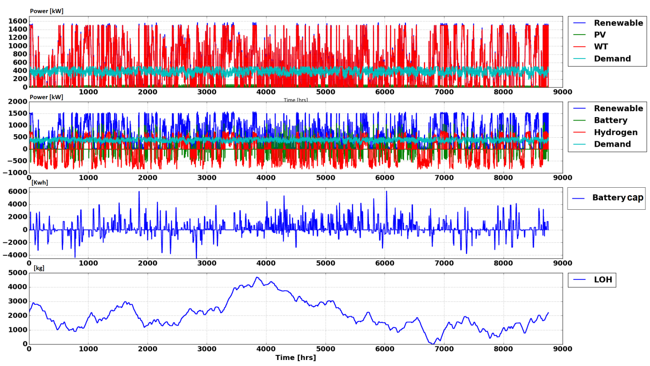
(a) Configuration A



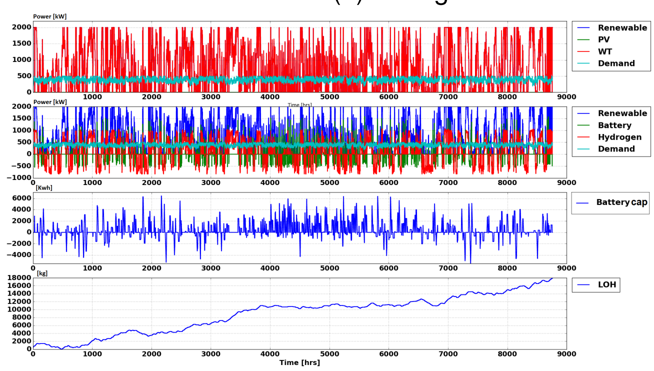
(b) Configuration B



(c) Configuration C



(d) Configuration D



(e) Configuration E

Figure 3.15: The results of the sizing strategy

- The battery is considerably less used: The previous case was a full PV which means zero energy production during the night so the batteries discharge to maintain the demand. In this case, wind turbine supplies the missing part of the data center demand..
- The level of hydrogen is more fluctuating and one cannot really differentiate the seasonality (winter and summer). This is also due to the hybridity of the system.

The last Figure 3.15e is a full wind turbine system. In this configuration, one can see that the system is over sized as the level of hydrogen increases over time and does not come back to the initial state at the end of the year. This is due to the energy production of the fourth wind turbine. Thus the case of full wind turbine cannot be used with this strategy.

3.5.4/ Applied performance results

3.5.4.1/ Cost Analysis

The total cost of each configuration obtained from the simulation is calculated using Equation (3.13). The results are illustrated in Figure 3.16 which shows the total cost of the different types of configuration per year.

Actually, Full PV means a configuration where the primary sources are only composed from PV, Hybrid 1 is a hybrid configuration composed from only one wind turbine and photovoltaic panels, Hybrid 2 is composed from 2 wind turbines and photovoltaic panels. Finally, Full WT is a configuration composed from only wind turbines. In Figure 3.16, one can identify that the configuration Full WT is the most expensive infrastructure over the years. It turn around the same prices, respectively 2500 kUSD and 3100 kUSD from 2004 till 2011.

The configuration Full PV is the second most expensive infrastructure overs the years. It costs around 2500 kUSD and 2700 kUSD depending of the executed year.

The hybrid configurations remains the cheapest infrastructure with some low differences; The Hybrid 1 and Hybrid 3 configurations are around 200k USD more costly than the Hybrid 2 configuration. Thus, in terms of cost, the Hybrid 2 configuration composed of 2 wind turbines and the rest of PV constitutes the best choice. Using this configuration, one can recognize the best year where the system is able to supply the data center with a minimal cost. This year is 2004 where the cost of the Hybrid 2 configuration is exactly equal to equal to $1,836,418.494$ USD and the worst year where the system is able to supply the data center with the maximal cost. This year is 2010 where the cost of the hybrid 2 system is 2,378,959.034 USD. .

3.5.4.2/ Reliability Analysis

The sizing strategy is made to turn the data center on stand alone and to be completely reliable. After simulation, the LPSP value for all the configurations obtained was equal to 0. It means that our system is completely reliable.

3.5.4.3/ Renewable energy penetration Analysis

As explained in the previous section, the REP analysis consists in the percentage of hours where the renewable energy supplies the data center without using the storage energy.

In Figure 3.17, the renewable energy penetration of each configuration (Full PV, Hybrid 1, Hybrid 2, etc. and Full WT) is illustrated in bar charts per year. The Full PV configuration turns out to be the worst in terms of autonomy. As there is no sun during the night, the Full PV configuration is directly related to the storage part. Thus, it is the configuration which loses the most energy in efficiency of the storage system. One can see that Hybrid 1, 2 and Full WT assure, in the majority of the cases, more than 50% of autonomy thanks to the availability of wind during the whole day. Nevertheless, the Hybrid 2 configuration is the configuration with the best REP value over the years except for the year 2010.

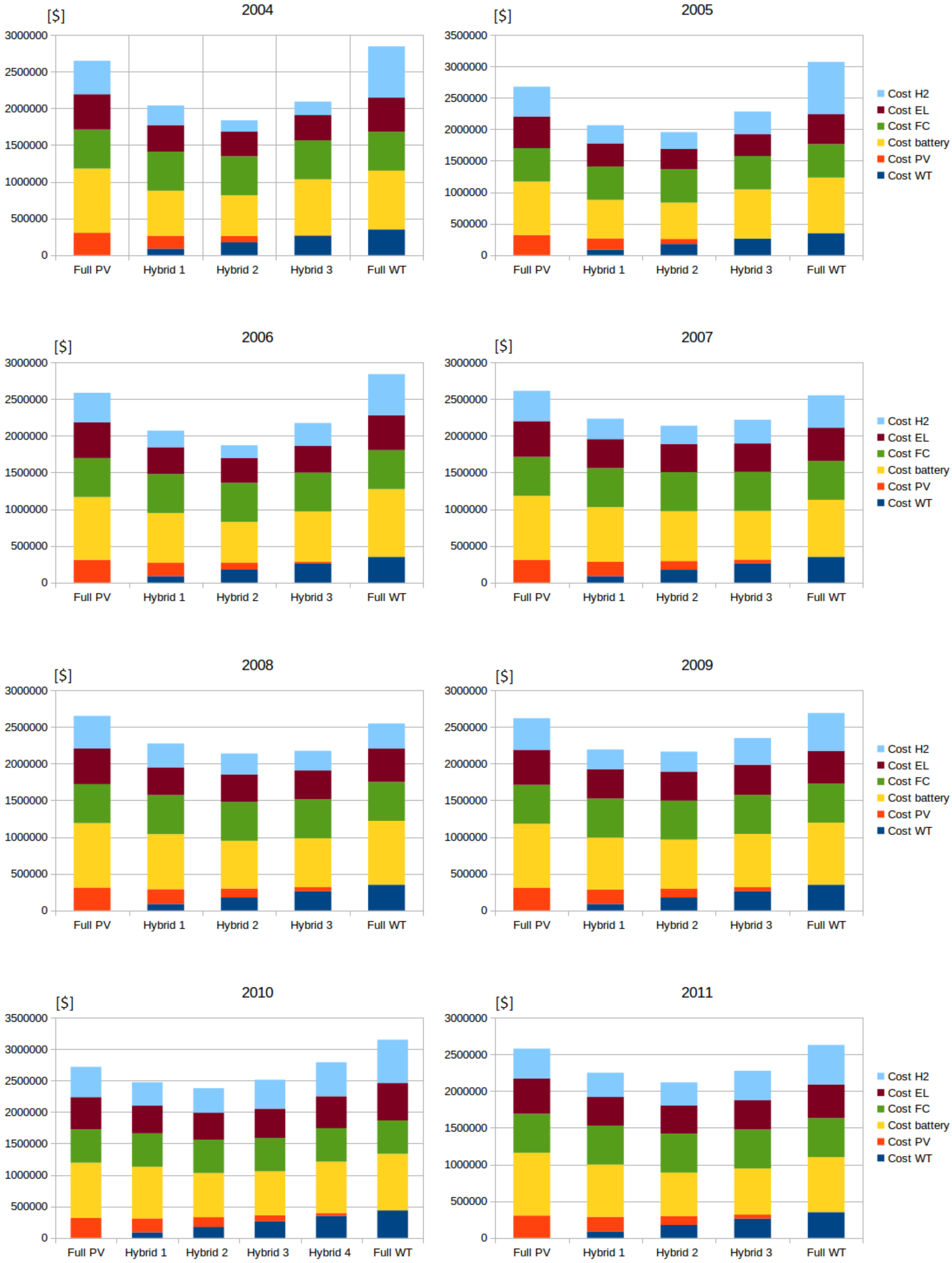


Figure 3.16: Cost analysis of the different configurations obtained for each year

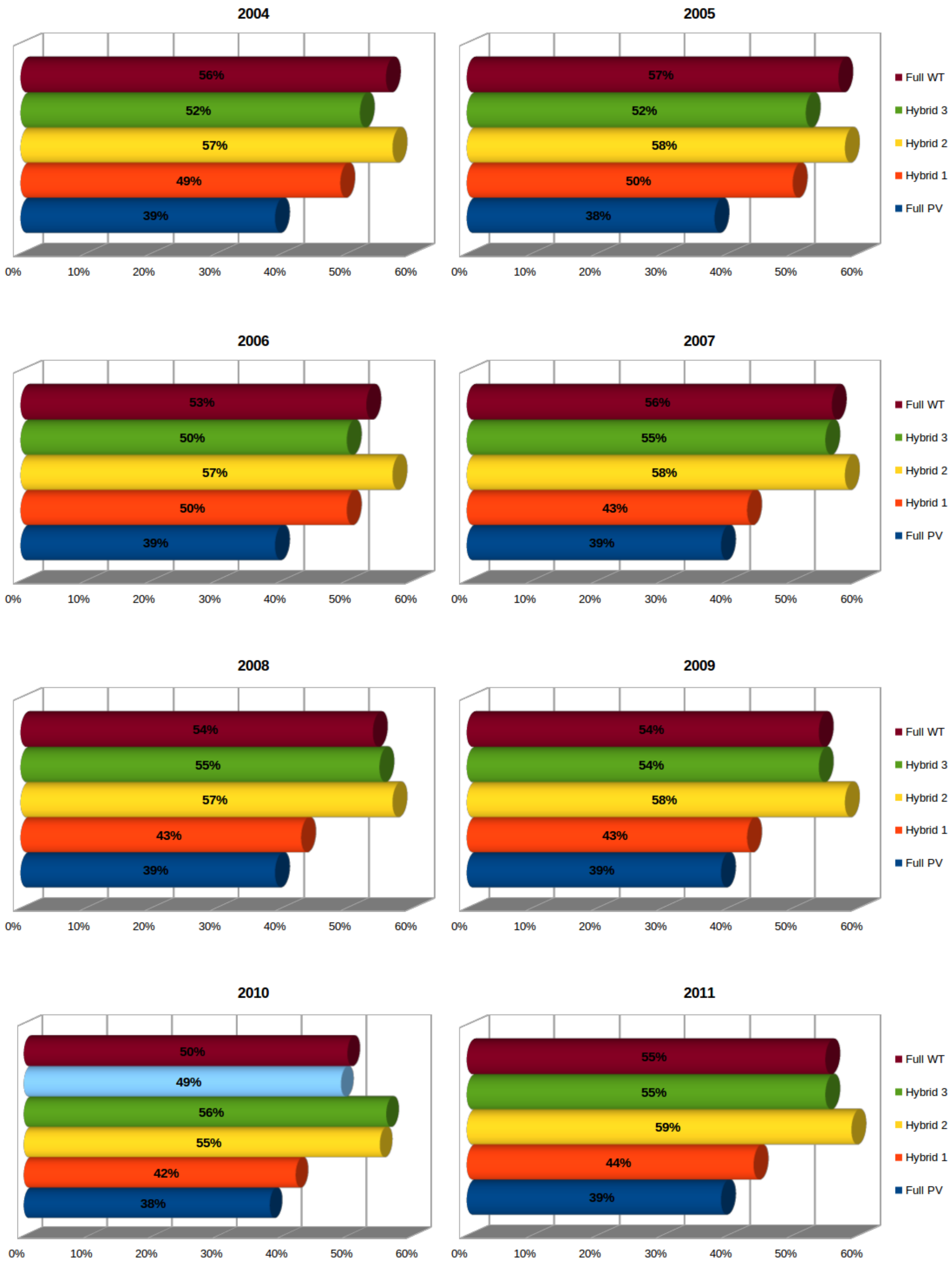


Figure 3.17: Autonomy analysis of the different configurations obtained for each year

3.5.5/ Summary of the results

The results obtained above suggest several interesting observations:

- All the simulations made for each configuration show an LPSP equal to 0: **The system is reliable for each configuration.**
- The simulations of the Full WT configurations does not correspond to the sizing strategy as the level of hydrogen increases at the end of the horizon. Second, their annual cost is really high compared to the other configurations.
- The Full PV configuration is quite costly for an autonomy quite low.
- Taking the renewable energy penetration metric into account, the best hybrid configuration is **Hybrid 2** composed of two wind turbines and the rest of solar panels as it has the best value over years.

⇒ The best configuration is the Hybrid 2 configuration to supply the data center.

Using this configuration (Hybrid 2), one can recognize first the best year where the system is able to supply the data center with a minimal cost. This year is 2004. Second, one can identify the worst year where the system is able to supply the data center with the maximal cost. This year is 2010. Using these two obtained sizing results, one can play them on the other years in order to see the storage variation.

3.6/ Playing sizing scenarios

The aim of this section is to play different sizings obtained of the configuration Hybrid 2 composed of 2 wind turbines and PV panels under other meteorological data of the other years in order to qualify the best sizing of the system. Three sizing values will be played in other scenarios:

- Best year: Supply the data center with minimal cost (2004)
- Worst year: Supply the data center with higher cost (2011)
- Average year: Calculate the average of the several sizing obtained for photovoltaic panels and wind turbines

3.6.1/ Best year

The sizing of the best year is composed of 2 wind turbine with a rated power of $500\,kW$ and $3791\,m^2$ with a unit power of $300\,kW$ of PV panels. This sizing is used under the meteorological data from 2004 to 2012 hour by hour. The results are illustrated in Figure 3.18. In this figure, the renewable energy produced by the wind turbines is in red, the energy produced by the PV panel in green, their sum in blue and the data center demand is in cian blue. One can also see the solicitation of the batteries during hours where the level comes back to the initial level of the battery. Finally, during the first year 2004, the level of hydrogen comes back to the initial level 0. During the other years, the level of hydrogen is quite discharging. Each peak actually corresponds to buying the needed quantity of hydrogen to restart over for the following year.

Thus, the best year sizing is not really an accurate sizing as it cannot handle the other years. Then this sizing scenario will not be chosen for our system.

3.6.2/ Worst year

The sizing of the best year is composed of 2 wind turbines and $6815m^2$ of PV panels. This sizing is used under the meteorological data from 2004 to 2012 hour by hour. The results are illustrated in Figure 3.19. In this figure, the renewable energy produced by the wind turbine is in red, the energy produced by the PV panels in green, their sum in blue and the data center demand is in cian blue.

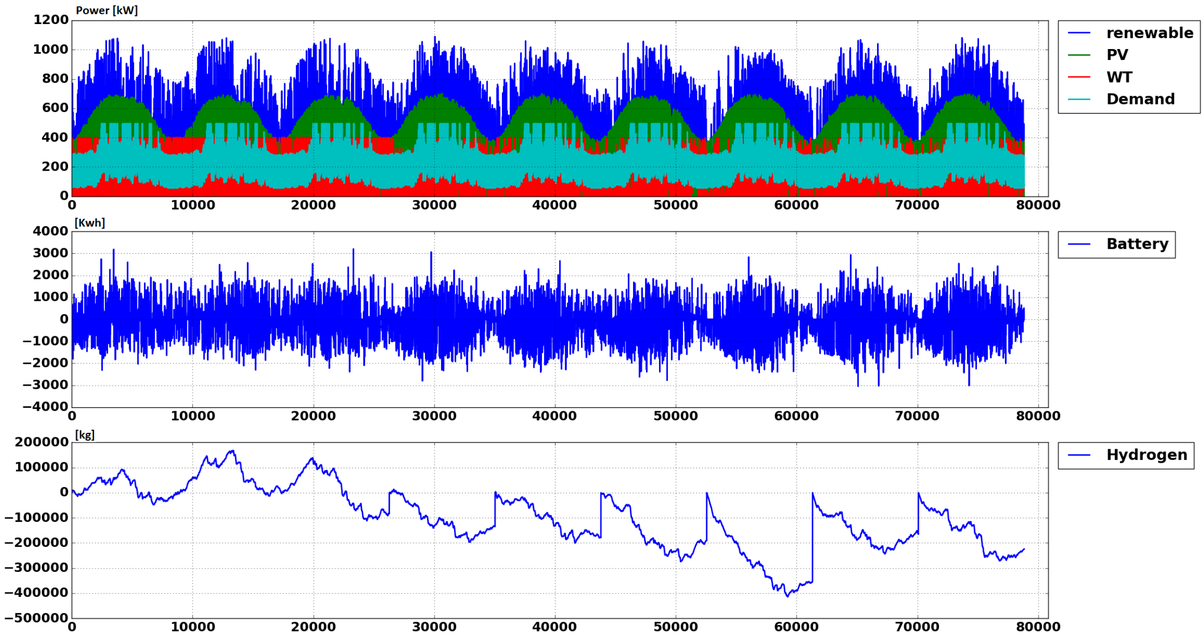


Figure 3.18: Scenario 1: Best year (2 WT & $3791\,m^2$ PV)

One can also see the solicitation of the batteries during hours where the level comes back to the initial level of the battery. Finally, during the first years, the level of hydrogen is quite increasing. The quantity overproduced of hydrogen is then sold at the end of each year. One can recognize the year 2010 on which the sizing was configured as the level of hydrogen comes back naturally to 0 at the end of the horizon.

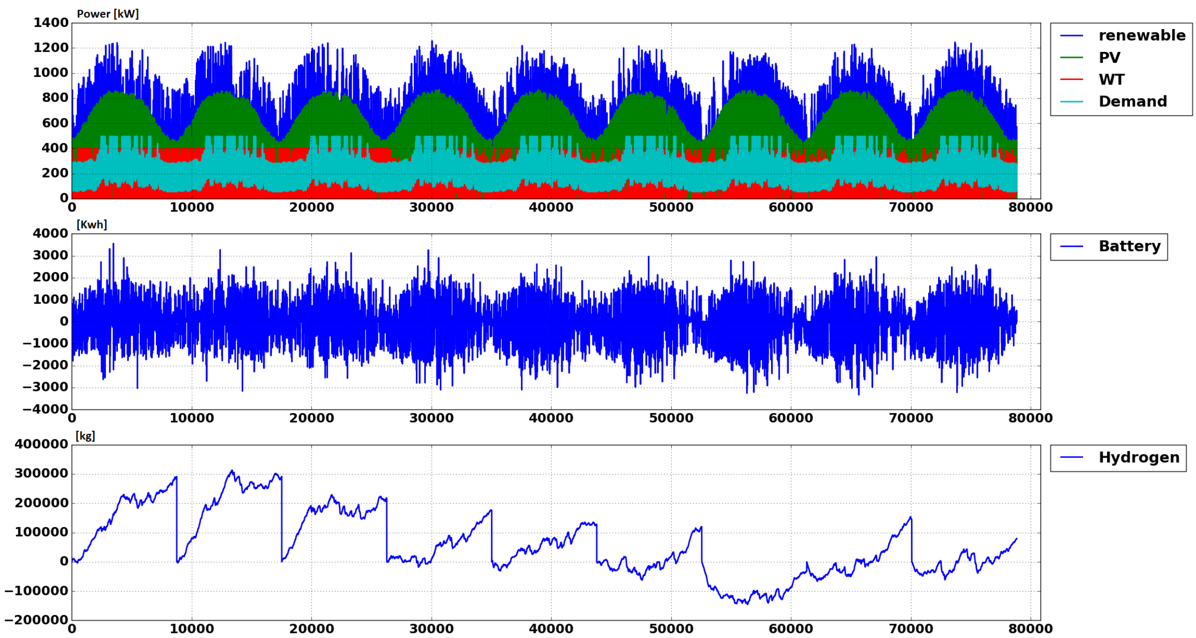


Figure 3.19: Scenario 2: Worst year (2WT & $6815m^2$ PV)

3.6.3/ Average year

The sizing of the average year was made on calculating the average value of the different PV areas and number of wind turbines obtained from sizing of each year from 2004 till 2012. Thus, this sizing is composed of two wind turbine and $4787m^2$ of PV panels. Figure 3.20 shows that the level of hydrogen is fluctuating over the years where in some years (2010 and 2011), there is need to buy hydrogen and in others (2004, 2005, 2006), there is need to sell it in order to answer to the demand of the data center.

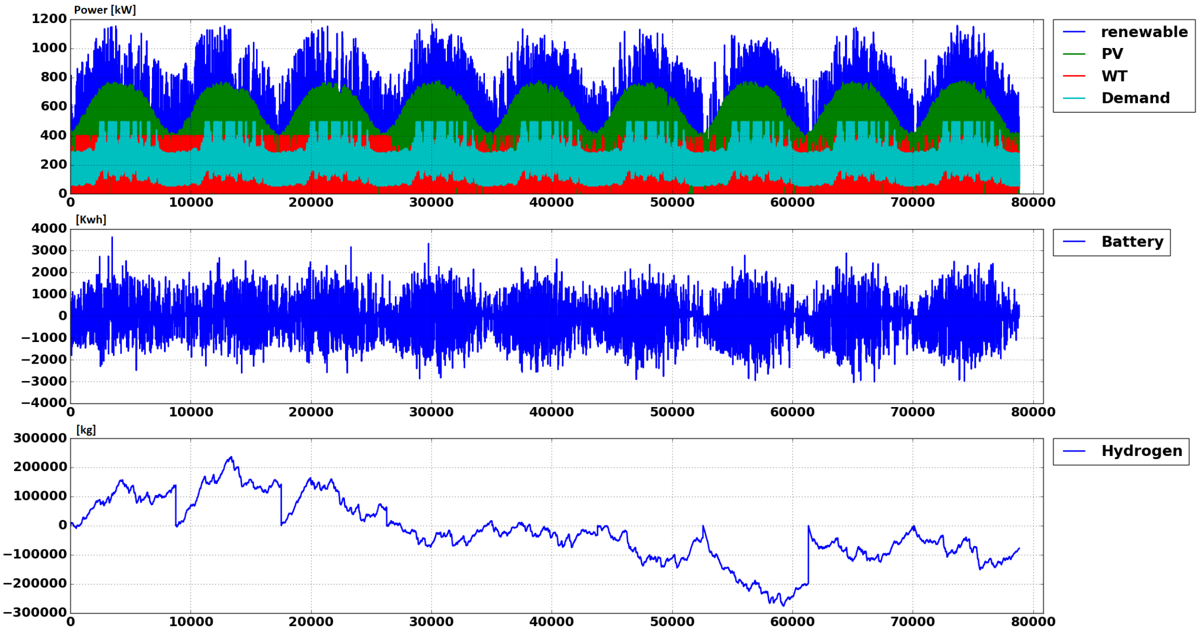


Figure 3.20: Scenario 3: Average year (2WT & 4787 PV)

3.7/ Synthesis

In this chapter, we presented a sizing strategy for a stand alone hybrid energy system composed of photovoltaic panels, wind turbines, batteries, fuel cells and electrolyzers. First, we presented the principle of our sizing strategy where the problem was taken backward. (determination of the power load such as the overproduction is equal to underproduction). Then, our final sizing strategy capable to respond to the data center demand is explained. The tool provided different results for each year with different configurations; Full PV, Hybrid 1 composed of 1 wind turbine and the rest of PV panels, Hybrid 2, etc, or Full WT. After introducing and analyzing some metrics such as the annual cost, the renewable energy penetration and the reliability, the best configuration chosen among all is Hybrid 2 as it is the best in quality/services. Once the configuration is fixed, we needed an exact sizing for our data center demand, thus we played different scenarios with different sizing such as the sizing of the best year, worst year and average year.

Chapter 4

Power Decision Module

This chapter aims to determine different optimal management policy for the hybrid renewable energy system in order to meet several objectives defined the DATAZERO project. This chapter starts by a definition of the power decision module, its architecture and its operating. The second part is dedicated to a synthesis of the hybrid renewable energy system models, presented in the precedent chapters (model (A)). These models are given in the form of a constraint satisfaction problem. Nevertheless, as this models are non linear, we proceed, in a second time, in the linearization of these constraints in order to construct a mixed integer linear optimization program (MILP). This chapter ends up by playing different scenarios of the mixed integer linear program in order to be able to deliver the power load needed by the negotiation module.

Contents

4.1/ Power Decision Module (PDM) overview

4.1.1/ PDM within DATAZERO

The Figure 4.1 gives an overview of the global structure of the DATAZERO system on which it is possible to see the different components that are connected to the middleware and the messages that can be exchanged between components. DATAZERO embeds several optimization problems as explained in the general introduction.

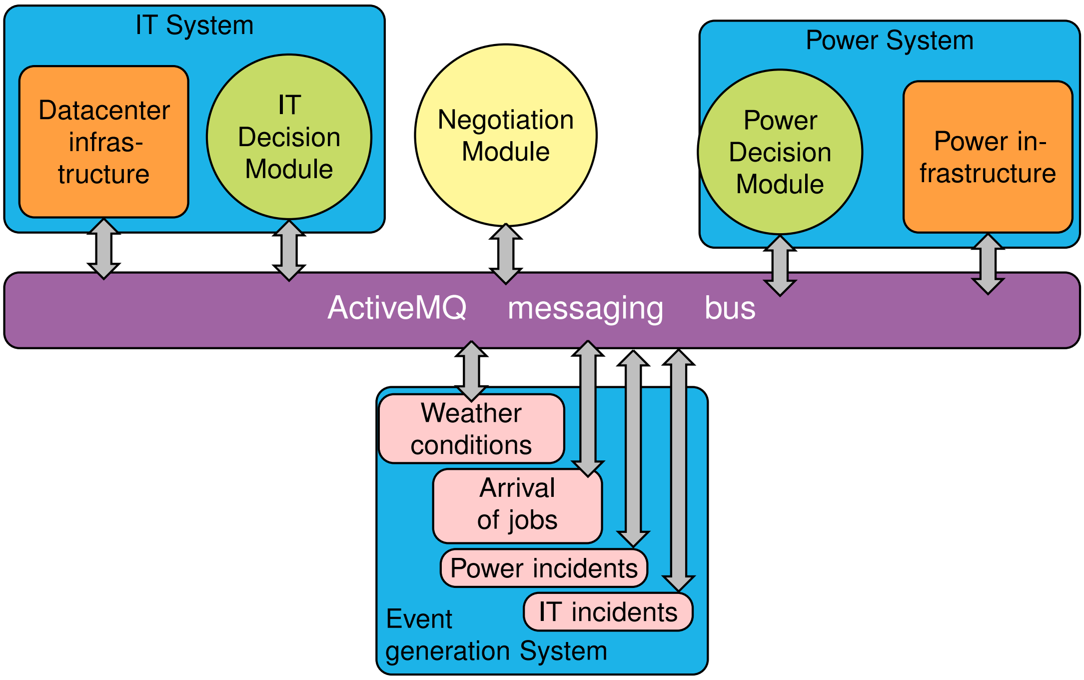


Figure 4.1: DATAZERO global structure

In fact, the IT and Power Decision Modules are the main places where optimization is necessary. On the one hand, in IT Decision Module, one has to deal with the allocation of the IT resources considering the available power while satisfying the required quality of service. On the other hand, in Power Decision Module, the required power should be produced using the available energy sources and storage, while considering the IT demands on the considered time horizon. As a consequence, various decision problems have to be solved. In this chapter, we propose to detail the problems related to the Power Decision Module (PDM). Each of them is described by several characteristics: known data or inputs of the problem, decision variables, constraints to satisfy, optimization objective(s) and decision time horizon. Moreover, as already explained in the general introduction and detailed in Figure 4.2, the negotiation module aims to guide both the decision optimization modules ITDM and PDM in order to converge to a power profile, first, close to the one requested by IT and second, feasible by the electrical part. In evidence, the IT Decision Module would prefer to execute jobs when that is better to ensure a Quality of Service (QoS). However, it may most likely not match with the best usage of renewable energy and storage devices. Then, it is mandatory to confront IT requests with the energy availability which is computed by the PDM.

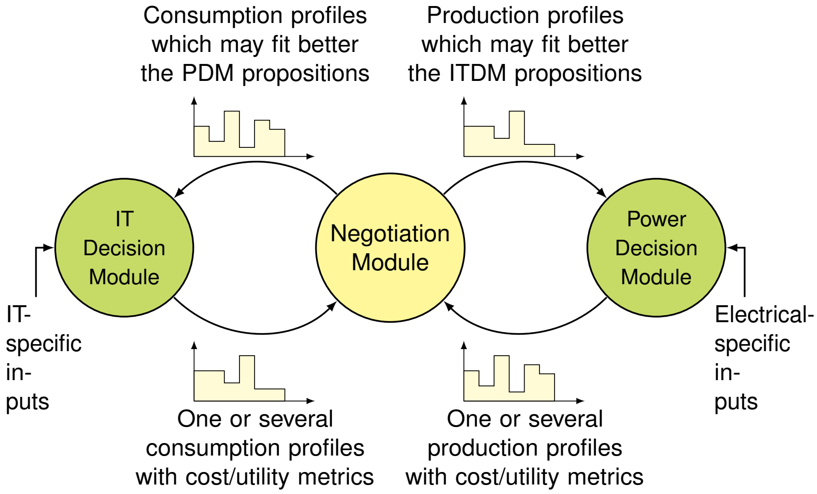


Figure 4.2: Illustration of the negotiation process, with the information exchanged between the decision modules and the negotiation module.

To summarize, the PDM is then considered as a black box used by the other modules of DATAZER to ensure an optimal management of the hybrid renewable energy sources able to meet the power demand of the data center sent by the negotiation module in each negotiation round. Once the negotiation is ended, the commitment of the hybrid system agreed on is then sent to the real infrastructure (Power Simulator or Power Hardware In the Loop in the figure 4.1) in order to be established for the planned horizon

4.1.2/ Power Decision module operations

The negotiation module forces both ITDM and PDM to design one or multiple power profiles and to send them for each negotiation round. In order to converge and to find the best commitment for both modules, the power profile sent must be different of the one sent before in the previous round of negotiation. Therefore, the power decision module addresses different optimization problems according to the purpose of the work, the constraints of the system as well as the particular objective. In the strategy proposed (see 4.3), the negotiation module (NM) starts by requesting different power profiles. PDM have to identify which source commitment can be proposed regarding the energy storage levels and the weather forecasting. In this part, the power decision module design two different profiles without having any information about the power demand of the data center. These profiles, once sent to the negotiation, are then used as a base by the latter to request against proposals. These two profiles are done following two different goals:

1. At the beginning, when a new negotiation is initiated (NegoRequestToPDM), the power decision module (PDM) receives the forecasted meteorological data, the electrical infrastructure and the level of the storage devices. Then, two expected power profiles are computed and sent as an answer to the negotiation module as follows:

    - Providing the maximum constant power profile while respecting all the constraints and the storage state and taking into account the weather forecasting during the chosen time horizon;

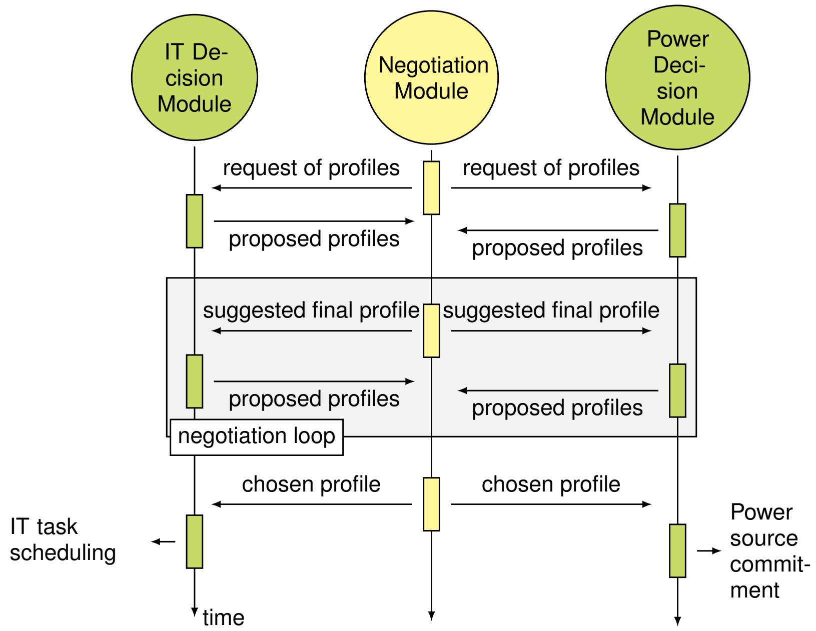


Figure 4.3: Illustration of the negotiation process, with the information exchanged between the decision modules and the negotiation module.

- Providing a non constant power profile where the goal consists in minimizing the loss due to the usage of electrical devices while respecting all the constraints and the storage level.

2. If the negotiation fails in finding a good compromise with the previously sent profiles, a second negotiation take place. In this round of the negotiation (part framed in gray in Figure 4.3), the negotiation module (NM) sends to PDM a power profile and asks to design a new profile that is close to the given one as far as possible (which means that the profile is computed while taking into account the renewable energy produced and respecting the state of storage). In this case, the PDM uses another optimized program with another goal in order to conceive this new eligible profile :

- Matching with the given load profile. The problem consists in maximizing the overall power production. Additional constraints force the power production for each period to be within an interval of values set by the suggested power profile sent by the NM, while respecting the constraints defined by the model and the storage state.

When a compromise is finally found, the negotiation sends a message that ends the negotiation round with the chosen power profile.

3. Obviously, if the proposed power profile meet the asked power proposal of NM then the negotiation is resolved. Then, the NM sends the chosen profile to both PDM and ITDM parts. The last optimization solver is requested to optimally find a commitment source by source. The source commitment obtained to implement optimally the negotiated power profile is then sent to the real power infrastructure by the PDM in order to start operating in the horizon planned. In this interaction, we assume that no external event occurs (new negotiation, or PSS change). Thus, this third problem follows another goal:

- Satisfying the load demand. The strategy proposed is to maximize the long-term storage by setting a target level to reach at the end of the horizon. This allows to ensure the seasonality variation of renewable energy production.

4. The PDM receives periodically the state of the platform to update its information or alarms from the power infrastructure. These communications are able to force a new negotiation round before the end of the current period because of an unexpected situation (not enough wind as expected, etc.). The new negotiation round takes updated weather forecasted data to propose expected/relaxed power profiles.

To conclude, the power decision module play an important role as a negociator with the negotiation module in order to find the best electrical source commitment and also, as a commander for the real electrical platform.

4.1.3/ Architecture of the PDM

In order to make this operations possible, the power decision module must be composed of several internal components which are responsible to listen the messages from the message bus (ActiveMQ), process the messages, propose possible power profiles with their utility, calculate electrical commitment, and interact with the real infrastructure (Power Hardware In the Loop – PHIL), the Power System Simulator (PSS) and the negotiation module (See Figure 4.4). Each component is described separately below, followed by an example of interaction with the Negotiation Module.

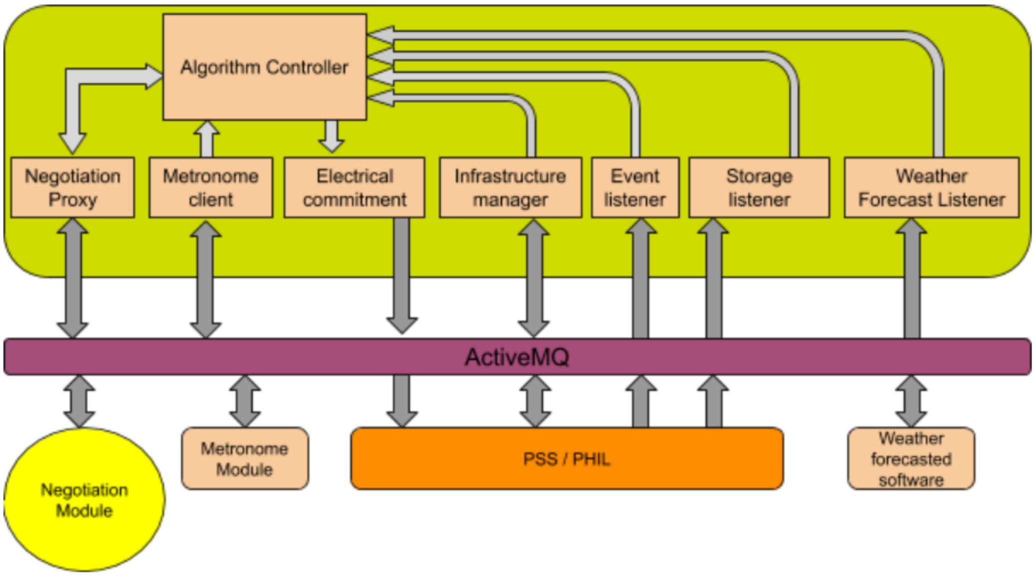


Figure 4.4: Power Decision Module architecture

- Weather Forecast Listener: receives the current and forecasted meteorological data and forwards these data to the Algorithm Controller to help it to compute power profiles and source commitment.
- Storage Listener: receives the level of each storage device (battery and hydrogen tank) from the power simulator (PSS) or power hardware in the loop (PHIL) before each negotiation.
- Event Listener: receives events that make the source commitment impossible. This implies a new negotiation round including the new configuration described within the event message.
- Infrastructure Manager: reads the initial infrastructure from a file made by the sizing module and sends it to the message bus (ActiveMQ) (to inform the PSS module). If the configuration of the electrical infrastructure is changing because of external events (wind turbine maintenance, etc.), the infrastructure is updated by listening the right topic of ActiveMQ. This information comes from the electrical part (PSS or PHIL).

- Electrical commitment: sends the power profile of each source device to the PSS or PHIL which has to follow the commitment once the negotiation ends.
- Metronome Client: stores, within its own event list, all events that are already scheduled. The Metronome Client sends to the Metronome Module (via ActiveMQ) the first upcoming event of the list. The Metronome Manager sends back the date of the next event to be globally scheduled within the DATAZERO platform. The Metronome Manager sends periodically its message until this event becomes the event to be scheduled.
- Negotiation Proxy: receives and parses the messages concerning the step of the negotiation round. The appropriate optimization algorithm is called within the Algorithm Controller. This proxy sends back to the Negotiation Module an asked power profile.
- Algorithm Controller: receives requests from the negotiation module. It is composed of different optimization solvers that aim at (1) creating and evaluating several profiles that fit as much as possible within every source state and the weather forecasting, (2) modifying and evaluating a proposed profile and (3) computing each source commitment for the next chosen horizon of time. The evaluation of the given profile helps the negotiation to find the appropriate compromise between the power demand from the IT part and the power offer from the Power System part.

The architecture will enable the processing of the power decision modules in order to achieve the explained goals. In order to better explain the optimized profiles sent in each negotiation round, the constraint used are explained in the next section.

4.2/ Constraints of the Electrical Part

In this section, the scientific contributions of the algorithm manager of the PDM is explained. It describes the usage limits of the different electrical devices (constraints). In this constraint system, one needs to define mathematically each goal needed to achieve for the negotiation process. To make the optimization process readable, a list of the notations used is summarized within Table 4.1.

4.2.1/ Power conservation

The average power produced ($Pwt_k + Ppv_k$) by the primary sources during the time step k can be for three purposes:

- Supply the electrolyzer (Pez_k)
- Charging the batteries (Pch_k)
- Power the data center demand ($Pload_k$)

In case there is not enough renewable average power to meet the power demand ($Pload_k$), additional electrical power is delivered by the fuel cells (Pfc_k) and batteries ($Pdch_k$). To satisfy the demand, Equation (4.1) should be fulfilled for each k with $k \in [\![0, K-1]\!]$:

$$Pload_k \leqslant Pwt_k + Ppv_k + (Pfc_k + Pdch_k - Pez_k - Pch_k) \times \eta_{inv} \tag{4.1}$$

where : η_{inv} is the efficiency of the inverter. This equation will be used as the power conservation equation in all the proposed problems defined in the following section. Moreover, while combining this equations and those characterizing the models (A) of each power sources defined in the second chapter (model of the photovoltaic panels, model of the wind turbine, etc.), one can define a constraint satisfaction problem.

Table 4.1: Nomenclature

Variable	Description
$\mathcal{H}$	A given time window $[h]$
Δt	Interval of time between two time steps $[h]$
k	Index for one time step within $\mathcal{H}$
K	Number of time steps within $\mathcal{H}$
The power of each sources is the mean power on the time step considered	
Ppv_k	PV power $[kW]$
Pwt_k	WT power $[kW]$
$SOCmax$	Maximal state of Charge (SOC) $[\%]$
$SOCmin$	Minimal SOC $[\%]$
SOC_k	SOC at instant $k\Delta t$ $\forall k \in [\![1, K]\!]$ $[\%]$
Pch_k	Power used to recharge battery $[kW]$
$Pdch_k$	Power discharged from battery $[kW]$
$Cbat$	Capacity of the battery $[kWh]$
$LOHtarget_D$	H_2 tank level targeted for a given day D
$LOHmax$	H_2 tank upper limit [%]
$LHVh_2$	H_2 lower heating value $[kWh.kg^{-1}]$
$HHVh_2$	H_2 higher heating value $[kWh.kg^{-1}]$
LOH_k	The level of H_2 at instant $k\Delta t$ $\forall k \in [\![1, K]\!]$ $[kg]$
H_2	Maximal tank capacity $[kg]$
$Pezmax$	Electrolyzer power upper limit [%]
$Pezmin$	Electrolyzer power lower limit [%]
Pez_k	Electrical Power supplied to electrolyzer at instant $k\Delta t$ $\forall k \in [\![0, K-1]\!]$ $[kW]$
Qez_k	Electrolyzer H_2 mass at instant $k\Delta t$ $\forall k \in [\![0, K-1]\!]$ $[kg]$
$Pfcmin$	Minimum power to operate fuel cell $[kW]$
$Pfcmax$	Maximum power delivered by fuel cell $[kW]$
Pfc_k	Power delivered by fuel cell at instant $k\Delta t$ $\forall k \in [\![0, K-1]\!]$ $[kW]$
Qfc_k	Fuel cell H_2 mass at instant $k\Delta t$ $\forall k \in [\![0, K-1]\!]$ $[kg]$
η_{ch}	Battery charge efficiency [%]
η_{dch}	Battery discharge efficiency [%]
σ	Battery self-discharge rate (constant) [%]
η_{ez}	Electrolyzer efficiency [%]
η_{fc}	Fuel Cell efficiency [%]
η_{inv}	Inverter efficiency (constant) [%]
x_k	Battery in use $(x_k = 1)$ or not $(x_k = 0)$
y_k	Electrolyzer in use $(y_k = 1)$ or not $(y_k = 0)$

4.2.2/ Constraint Satisfaction Problem

The resulting model can be viewed as a Constraint Satisfaction Problem (CSP). The CSP is defined as follows (with $k \in [\![0, K-1]\!]$ or $k \in [\![0, K]\!]$ for SOC_k and LOH_k variables):

$$Pload_k \leqslant Pwt_k + Ppv_k + (Pfc_k + Pdch_k)\eta_{inv} - (Pez_k + Pch_k)\eta_{inv} \tag{4.1}$$

$$SOC_k = \min\{SOC_{k-1} \times (1-\sigma) + \frac{Pch_{k-1} \times \eta_{ch} \times \Delta t}{Cbat}, SOCmax\} \quad \text{if } Pch_{k-1} > 0 \tag{4.2}$$

$$SOC_k = \max\{SOC_{k-1} \times (1-\sigma) - \frac{Pdch_{k-1}}{\eta_{dch} Cbat} \times \Delta t, SOCmin\} \quad \text{if } Pdch_{k-1} > 0 \tag{4.3}$$

$$Pez_k \Delta t = HHVh_2 \times Qez_k / \eta_{ez} \tag{4.4}$$

$$Pfc_k \Delta t = LHVh_2 \times Qfc_k \times \eta_{fc} \tag{4.5}$$

$$LOH_k = LOH_{k-1} + Qez_{k-1} - Qfc_{k-1}/\eta_{tank} \tag{4.6}$$

Bounds: (4.7)

$$Pfc_k \leqslant Pfcmax \tag{4.8}$$

$$Pezmin \leqslant Pez_k \leqslant Pezmax \tag{4.9}$$

$$SOCmin \leqslant SOC_k \leqslant SOCmax \tag{4.10}$$

$$0 \leqslant LOH_k \leqslant LOHmax \tag{4.11}$$

The batteries are used for short term storage. The state of charge of the batteries is calculated for each k with $k \in [\![1, K]\!]$ with respect to the previous state of charge of the battery SOC_{k-1}, the self-discharge rate σ, the charging and discharging power Pch_{k-1}, $Pdch_{k-1}$, the charging and discharging efficiency η_{ch}, η_{dch} as follows:

$$SOC_k = SOC_{k-1}(1-\sigma) + \frac{Pch_{k-1}\Delta t \times \eta_{ch} - Pdch_{k-1}\Delta t/\eta_{dch}}{Cbat} \tag{4.12}$$

As the battery storage system (BSS) cannot charge and discharge simultaneously, the value of SOC_k is bounded by the minimum amount of $SOCmin$ and maximum energy $SOCmax$. So $\forall k \in [\![1, K]\!]$ with $SOC_0 = SOC_{init}$ the state of charge of the battery is calculated as in the equations (4.2) and (4.3).

Equation (4.4) is the calculation of the operating power of the electrolyzer Pez_k depends of the hydrogen mass Qez_k in (kg), and is bounded by $Pezmin$, $Pezmax$ the operating range of the electrolyzer. This relation is defined for each time step k ($k \in [\![0, K-1]\!]$), with η_{ez} the electrolyzer efficiency and $HHVh_2$ the hydrogen higher heating value.

Equation (4.5) is the computing of the output power of the fuel cell Pfc_k at any time during each period k ($k \in [\![0, K-1]\!]$). It depends of the hydrogen mass Qfc_k and η_{fc} its efficiency and $LHVh_2$ the low heating value of hydrogen.

The level of hydrogen produced by the electrolyzer and consumed by the fuel cell in the tank is determined at any time during each period k with $k \in [\![1, K]\!]$ as in equation (4.6).

Nevertheless, almost all used equations are not linear. Then, in a second step, this model will be linearized in the next section to be solved using an efficient solver as Gurobi. This allows to propose solutions optimally computed by the Power module of the project: (i) identification of the profile that maximizes the power produced by the electrical sources used , (ii) computation of an optimal commitment around a target profile considering a given relax factor (which will be defined later). , and (iii) computation of an optimal commitment of the electrical sources constrained by the load profile.

4.2.3/ Context of the resolution

In addition, common sense rules have to be considered to allow the use of power components without any time restriction if both following constraints are respected:

- When a FC starts working, it is only used to satisfy the demand and not to charge batteries.

- When batteries start discharging, they are only used to satisfy the demand and not to produce hydrogen.

These common sense rules are fixed in order to avoid unnecessary power conversions that would increase losses. In the next section, we propose to transform the constraint satisfaction problem in a linear constraint satisfaction problem.

4.2.4/ Linearization of the constraints

In this section, in order to be solved, the proposed constraint satisfaction problem is linearized. Each constraint of the model is transformed to obtain a linear relation.

4.2.4.1/ The battery state of charge

The state of charge of the battery is limited by the SOC range ($SOCmin$ and $SOCmax$) with $k \in [\![1, K]\!]$ and depends both on the charge and discharge phases:

$$\begin{cases} SOCmin \leqslant SOC_k \leqslant SOCmax \\ SOC_k = SOC_{k-1}(1-\sigma) + \frac{Pch_{k-1}\Delta t \times \eta_{ch} - Pdch_{k-1}\Delta t/\eta_{dch}}{Cbat} \end{cases} \tag{4.13}$$

Because the both charge and discharge phases are concurrent phases, we introduce the set of binary variables x_k, and two sets of variables Pch'_k, $Pdch'_k$ with $k \in [\![0, K-1]\!]$ such as :

$$\begin{cases} Pch_k = x_k \times Pch'_k \\ Pdch_k = (1 - x_k) \times Pdch'_k \end{cases} \tag{4.14}$$

where $x_k = 0$ means that the battery is discharging and $x_k = 1$ means that the battery is charging. In this way, the battery can not charge and discharge at the same time. Then we propose to rewrite the previous state of charge equations as the following:

$$\begin{cases} SOCmin \leqslant SOC_k \leqslant SOCmax \\ SOC_k = SOC_{k-1}(1-\sigma) + \frac{x_k \times Pch'_{k-1} \times \Delta t \times \eta_{ch} - (1-x_k) \times Pdch'_{k-1}\Delta t/\eta_{dch}}{Cbat} \end{cases} \tag{4.15}$$

Due to these substitutions, we add new constraints to linearize (4.15):

$$\begin{cases} 0 \leqslant Pch_k \leqslant Pchmax \\ Pch_k \geqslant 0 \\ Pch_k \leqslant x_k Pchmax \\ Pch_k \leqslant Pch'_k \\ Pch_k \geqslant Pch'_k - (1 - x_k)Pchmax \end{cases} \tag{4.16}$$

$$\begin{cases} 0 \leqslant Pdch_k \leqslant Pdchmax \\ Pdch_k \geqslant 0 \\ Pdch_k \leqslant (1 - x_k)Pdchmax \\ Pdch_k \leqslant Pdch'_k \\ Pdch_k \geqslant Pdch'_k - x_k Pdchmax \end{cases} \tag{4.17}$$

Because Equation (4.14) is not written within the final linear program, the fact that the battery cannot charge and discharge is introduced by other new equations:

$$\begin{cases} Pch_k \leqslant x_k \times Pchmax \quad \text{with} \quad k \in [\![0, K-1]\!] \\ Pdch_k \leqslant (1-x_k) \times Pdchmax \quad \text{with} \quad k \in [\![0, K-1]\!] \\ x_k \in \{0, 1\} \end{cases} \tag{4.18}$$

where $Pchmax$ and $Pdchmax$ are the respective upperbound of Pch_k and $Pdch_k$.

4.2.4.1.1/ Battery operation: The battery is supposed to balance underproduction and overproduction during a period of time H. For instance if $\mathcal{H} = 24\,\text{h}$ with $\Delta t = 1\,\text{h}$ and $K = 24$, batteries are used for the fluctuations between day and night. It means that the production is smoothed over the day. To make it possible, the level of the state of charge of the battery should be greater or equal than its value at the beginning of each period $\mathcal{H}$. This operation is explained by the following to fix values before the resolution of the program if $\mathcal{H}$ is larger than $24\,\text{h}$:

$$\begin{cases} SOC_K = SOC_0 = SOCinit \text{ if } K < 24 \\ SOC_k = SOC_0 = SOCinit \text{ if } K \geqslant 24 \text{ and } k\%24 = 0 \ (k \in [\![0, K]\!]) \end{cases} \tag{4.19}$$

with $SOCinit$ the right level of charge to make possible the daily power compensation from hours where an extra renewable production is existing to hours this production is lacking. That is why this level has to come back to the same level each day. The difference between days is compensated by using long term storage.

4.2.4.2/ Electrolyzer

When the electrolyzer operates, it must respect the power range, otherwise it will not work and its power will be equal to zero. The following bounds express these cases ($\forall k \in [\![0, K-1]\!]$):

$$\begin{cases} Pezmin \leqslant Pez_k \leqslant Pezmax \\ Pez_k \geqslant 0 \end{cases} \tag{4.20}$$

It is then necessary to introduce a set of binary variables y_k to model the two alternatives where (with $k \in [\![0, K-1]\!]$):

- $y_k = 1$ means that the electrolyzer is in use with $Pezmin \leqslant Pez_k \leqslant Pezmax$ during period k;
- $y_k = 0$ means that the electrolyzer is stopped during period k.

By adding this binary variables, the equations (4.20) become as follows:

$$\begin{cases} (1-y_k) \times Pez_k = 0 \\ y_k \times Pez_k \leqslant Pezmax \\ y_k \times Pez_k \geqslant y_k \times Pezmin \\ y_k \in \{0, 1\} \end{cases} \tag{4.21}$$

Due to the non linearity of the set of equations (4.21), we propose to add a new variable substitutions Pez'_k such as:

$$Pez_k = y_k \times Pez'_k \tag{4.22}$$

then we add the following equations to linearize (4.21):

$$\begin{cases} Pez_k \leqslant Pez'_k \\ Pez_k \geqslant 0 \\ Pez_k \leqslant y_k \times Pezmax \\ Pez_k \geqslant Pez'_k - (1 - y_k)Pezmax \\ 0 \leqslant Pez'_k \leqslant Pezmax \\ Pez_k \geqslant y_k \times Pezmin \\ y_k \in \{0, 1\} \end{cases} \tag{4.23}$$

4.2.4.3/ Hydrogen tank level

As mentioned before for the battery, it is necessary to manage Qfc_k and Qez_k in order to control the hydrogen mass during the k^th time step Δt in the hydrogen tank as mentioned in Equation (4.6). Equations (4.4), (4.5) and (4.6) can be written by considering the new definition of Qez_k and Qfc_k for $k \in [\![0, K]\!]$:

$$\begin{cases} Pez_k \Delta t = HHVh_2 \times Qez_k/\eta_{ez} \\ Pfc_k \Delta t = LHVh_2 \times Qfc_k \times \eta_{fc} \\ LOH_k = LOH_{k-1} + Qez_{k-1} - Qfc_{k-1}/\eta_{tank} \end{cases} \tag{4.24}$$

Therefore, additionally to the two new sets of variables Qez'_k and Qfc'_k, we introduce an other set of binary variables z_k to model the two alternatives where the fuel cells are in use or not (with $k \in [\![0, K-1]\!]$). As before and for the same reasons, let be the following variable substitutions:

$$\begin{cases} Qez_k = z_k \times Qez'_k \\ Qfc_k = (1 - z_k) \times Qfc'_k \end{cases} \tag{4.25}$$

where:

- $z_k = 1$ means that $Qez_k = 0$
- $z_k = 0$ means that $Qfc_k = 0$

then we add the following set of equations :

$$\begin{cases} Qez_k \leqslant Qez'_k \\ Qez_k \geqslant 0 \\ Qez_k \leqslant z_k Qezmax \\ Qez_k \geqslant Qez'_k - (1 - z_k)Qezmax \\ 0 \leqslant Qez_k \leqslant Qezmax \end{cases} \tag{4.26}$$

$$\begin{cases} Qfc_k \leqslant Qfc'_k \\ Qfc_k \geqslant 0 \\ Qfc_k \leqslant (1 - z_k) \times Qfcmax \\ Qfc_k \geqslant Qfc'_k - z_k \times Qfcmax \\ 0 \leqslant Qfc_k \leqslant Qfcmax \end{cases} \tag{4.27}$$

In the rest of the document, we use these constraints as a basic model for the different applications of the PDM.

4.2.5/ Additional linear constraints

As indicated in Section 4.2.3, let us recall the common sense usage rules.

4.2.5.1/ Mutual exclusion between hydrogen and battery

The valid usage cases are:

1. The battery is in charge ($x_k = 1$) while electrolyzer can be in use ($y_k = 0$ or 1) and obviously fuel cell is stopped ($z_k = 0$).
2. The battery is discharging ($x_k = 0$) then fuel cell can be in use ($z_k = 0$ or 1) and the electrolyzer has to be stopped ($y_k = 0$).
3. The electrolyzer is in use ($y_k = 1$) with the battery in charge or not ($x_k = 0$ or 1), then the fuel cell is stopped ($z_k = 0$).
4. Start and stop of the fuel cell and the electrolyzer are allowed in this problem.

The following constraints represent the solution to these requirements:

$$\begin{cases} (1 - x_k) \times y_k = 0 \\ x_k \times y_k = y_k + x_k \times z_k \end{cases} \tag{4.28}$$

Due to the non linearity of the set of equations (4.28), we propose to add a new set of variables u_k and v_k (with $k \in [\![0, K-1]\!]$) that express the mutual exclusion between hydrogen and battery where:

- $u_k = x_k \times y_k$
- $v_k = x_k \times z_k$

then we add the following set of equations :

$$\begin{cases} u_k \leqslant x_k \\ u_k \leqslant y_k \\ 1 - x_k - y_k + u \geqslant 0 \\ u_k \geqslant 0 \end{cases} \tag{4.29}$$

$$\begin{cases} v_k \leqslant x_k \\ v_k \leqslant z_k \\ 1 - x_k - z_k + v \geqslant 0 \\ v_k \geqslant 0 \end{cases} \tag{4.30}$$

The obtained mixed integer linear program will be used in the following section as a basic model for the different optimization programs designed for the power decision module. Using this model, one can find out how engaging the electrical sources in order provide enough power to the data center following the defined goal and while respecting the constraints.

4.3/ Optimization programs obtained

In this section, the Power Decision Module is the used in order to find solutions of the optimization problems that will be explained below according to the negotiation state. These optimization problems are solved using linear programming which is defined by a set of equations: an objective function and some constraints. Following the stage of the negotiation, the PDM should be able to provide the good commitment of the power sources. Thus, as detailed in section 4.1.2, the PDM is able to design different profiles in function of the goal it has to ensure.

The problem takes as input the renewable power produced under weather conditions $Prenew_k = Ppv_k + Pwt_k$, power demand $Pload_k$ and other information of sizing like number of wind turbines, PV panels, energy storage capacity and efficiency of any component, and generates a a power commitment of each time slot for supply-side source and energy storage usage, according to the availability of renewable power.

The optimization problems explained in the following paragraphs are programmed under python and solved using GUROBI solver. The same sizing and the same weather conditions are applied to all the programs for a better illustration of the difference of results.

4.3.1/ Providing a constant power profile

In this section, the aim is to obtain a maximum power production $Pprod$ from the electrical sources for the whole horizon $\mathcal{H}$, taking into account the efficiency of each storage part used. Then the PDM will take as input the meteorological data to determine the primary production, the storage situation and the target set. As an output, the PDM must send the maximum constant profile obtained and the storage situation to the negotiation module.

To make this possible, some modifications will be set for the mixed integer linear program and a new objective function and new constraints are added. To simplify it, the new variables added are explained in Table 4.2.

Table 4.2: Notations

Variable	Description
$Pprod = Pprod_k \quad \forall k$	The constant power profile $[kW]$
$LOHtarget_D$	A level of hydrogen defined by the sizing needed to be reached by the level of hydrogen LOH_K at the end of the horizon $\mathcal{H}$ $[kg]$

4.3.1.1/ Objective function

The objective function consists of maximizing the constant power profile obtained from the source commitment $Pprod$ in order to propose a profile in the first part to the negotiation. The problem is thus mathematically translated to:

$$\text{maximize} \quad Pprod \tag{4.31}$$

4.3.1.2/ Additional constraints

Some constraints are added or modified in order to consider common sense usage rules: consequently, Equation (4.1) is modified in this program as (4.32). Another constraint (4.33) is added to ensure that the computed hydrogen level at the end of the period of simulation reached the target defined by the sizing. In fact, while sizing, the usage of the storage system were played during the horizon $\mathcal{H}$. Thus, the level of hydrogen defined, in the sizing, is used as the hydrogen target for our simulations in order to support the considered horizon $\mathcal{H}$. It means that the computed level of hydrogen of the PDM is compared, at the end of the horizon, to the one defined by the sizing. This allows the storage system to handle the next periods $\forall k \in [\![0, K-1]\!]$:

$$Pprod \leqslant Pwt_k + Ppv_k + (Pfc_k + Pdch_k - Pez_k - Pch_k)\eta_{inv} \quad (4.32)$$
$$LOHtarget_D \leqslant LOH_K \quad (4.33)$$

The hydrogen target level $LOHtarget_D$ is supposed to be computed based on long term weather forecasting and IT load consumption for one year. D represents the day number within the current year and matches with the time horizon $K \times \Delta t$.

4.3.1.3/ Obtained model

The constraints explained in Section 4.2.4 are used for each version of the PDM as the basic constraints. Then, the linear program able to produce a constant maximum profile using the storage system composed of batteries and hydrogen system is then written as mentioned in the equation 4.34.

The obtained model used to provide a constant power profile as a first proposition for the negotiation is in the Equation (4.34).

4.3.1.4/ Results

In this linear program, we suppose that we are in the first stage of the negotiation which explains the absence of load in Figure 4.5. Actually, in that step of the negotiation, the PDM is supposed to provide the negotiation with profiles without getting data about the load. The same power architecture as in the previous section is used.

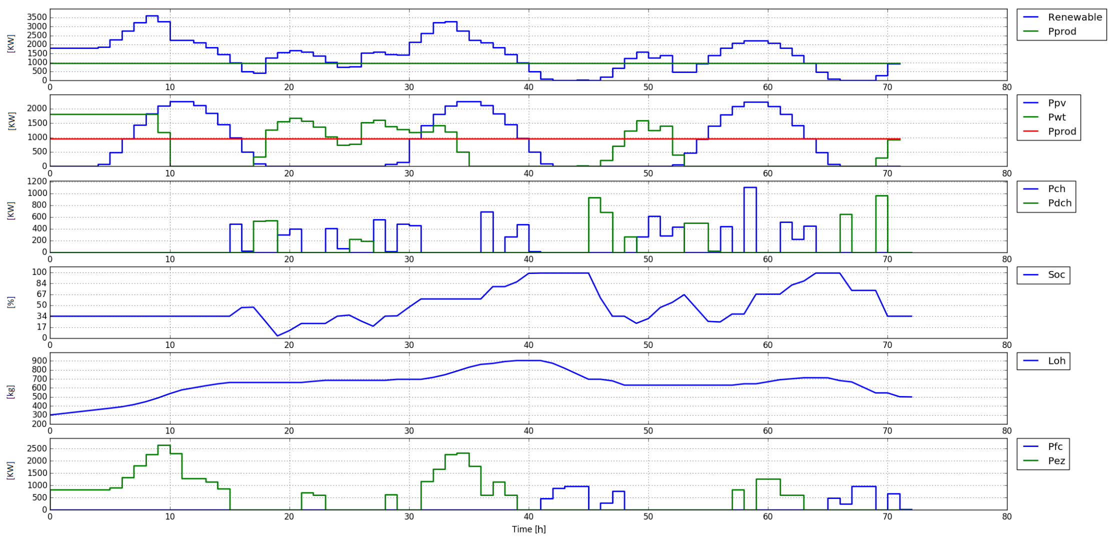


Figure 4.5: Constant Power Profile management

Results presented in Figure 4.5 are obtained by applying this linear program. The solution gives here a maximum constant power that can be delivered equals to 1 MW. As can be seen in the figure, the battery is responsible of the daily smoothing (day/night alternation). Thanks to the constraint (4.33), the level of hydrogen is even more than targeted at the end of the horizon H ($LOHtarget_D = 300kg$) to be able to assist the next time horizon.

$$
\left\{
\begin{array}{rcl}
 & & \text{maximize } Pprod \\
s.t.: & & \\
Pprod & \leqslant & Pwt_k + Ppv_k + (Pfc_k + Pdch_k)\eta_{inv} - (Pez_k + Pch_k)\eta_{inv} \\
SOC_k & = & SOC_{k-1}(1-\sigma) + \frac{\eta_{ch}Pch_{k-1}\Delta t - Pdch_{k-1}\Delta t/\eta_{dch}}{Cbat} \\
LOH_K & \geqslant & LOHtarget_D \\
Pload_k & \leqslant & Pwt_k + Ppv_k + (Pfc_k + Pdch_k)\eta_{inv} - (Pez_k + Pch_k)\eta_{inv} \\
Pez_k & = & HHVh_2 \times Qez_k/\eta_{ez}/\Delta t \\
Pfc_k & = & LHVh_2 \times Qfc_k \times \eta_{fc}/\Delta t \\
LOH_k & = & LOH_{k-1} + Qez_{k-1} - Qfc_{k-1}/\eta_{tank} \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \geqslant & 0 \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \leqslant & Pch'_k \\
Pch_k & \geqslant & Pch'_k - (1-x_k)Pchmax \\
Pdch_k & \leqslant & (1-x_k) \times Pdchmax \\
Pdch_k & \geqslant & 0 \\
Pdch_k & \leqslant & (1-x_k)Pdchmax \\
Pdch_k & \leqslant & Pdch'_k \\
Pdch_k & \geqslant & Pdch'_k - x_kPdchmax \\
Pez_k & \leqslant & Pez'_k \\
Pez_k & \geqslant & 0 \\
Pez_k & \leqslant & y_k \times Pezmax \\
Pez_k & \geqslant & Pez'_k - (1-y_k)Pezmax \\
0 & \leqslant & Pez'_k \leqslant Pezmax \\
Pez_k & \geqslant & y_k \times Pezmin \\
Qez_k & \leqslant & Qez'_k \\
Qez_k & \geqslant & 0 \\
Qez_k & \leqslant & z_k \times Qezmax \\
Qez_k & \geqslant & Qez'_k - (1-z_k)Qezmax \\
0 & \leqslant & Qez'_k \leqslant Qezmax \\
Qfc_k & \leqslant & Qfc'_k \\
Qfc_k & \geqslant & 0 \\
Qfc_k & \leqslant & (1-z_k) \times Qfcmax \\
Qfc_k & \geqslant & Qfc'_k - z_k \times Qfcmax \\
0 & \leqslant & Qfc'_k \leqslant Qfcmax
\end{array}
\right.
\quad
\left\{
\begin{array}{ll}
u_k \leqslant x_k & \\
u_k \leqslant y_k & \\
0 \leqslant 1 - x_k - y_k + u & \\
u_k \geqslant 0 & \\
v_k \leqslant x_k & \\
v_k \leqslant z_k & \\
0 \leqslant 1 - x_k - z_k + v & \\
v_k \geqslant 0 & \\
Bounds: & \\
\forall k \in [\![0, K-1]\!] & Pfc_k \leqslant Pfcmax \\
\forall k \in [\![0, K-1]\!] & Pez_k \geqslant Pezmin \\
\forall k \in [\![0, K]\!] & SOCmin \leqslant SOC_k \leqslant SOCmax \\
\forall k \in [\![0, K]\!] & 0 \leqslant LOH_k \leqslant LOH_{max} \\
\forall k \in [\![0, K-1]\!] & Pch_k, Pdch_k, Pez_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Pch'_k, Pdch'_k, Pez'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Qez_k, Qfc_k, Qez'_k, Qfc'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & x_k, y_k, z_k \in \{0, 1\}
\end{array}
\right.
\tag{4.34}
$$

4.3.2/ Providing a non constant power profile

In this section, the aim is to obtain a non constant power production profile $Pprod_k$ ($k \in [\![0, K-1]\!]$) from the electrical sources used, taking into account the efficiency of each storage part used. Then the PDM takes the meteorological data as an input to determine the primary production, the storage levels, etc., along the horizon $\mathcal{H}$.

As an output, the PDM must send the maximum power profile, and the storage situation to the negotiation module.

To make this possible, some modifications are set to the linear program obtained in the previous section. To simplify it, the variable $Pprod$ is no longer equal to a constant value during the horizon $\mathcal{H}$ but is changes following each step k in the horizon.

4.3.2.1/ Objective function

For this strategy, the objective function consists of maximizing the power profile obtained from the source commitment $Pprod_k$ ($k \in [\![0, K-1]\!]$) in order to propose a profile in the first part to the negotiation. The problem is thus mathematically translated to:

$$\text{maximize} \quad \sum_{k=0}^{K-1} Pprod_k \tag{4.35}$$

4.3.2.2/ Added constraints

As in the constant power case, some constraints are added or modified to consider common sense usage rules then Equation (4.1) is modified in this program as (4.36) ($\forall k \in [\![0, K-1]\!]$):

$$Pprod_k \leqslant Pwt_k + Ppv_k + (Pfc_k + Pdch_k - Pez_k + Pch_k)\eta_{inv} \tag{4.36}$$

The same constraint (4.33) is added to ensure the target level of hydrogen is reached at the end of the period of simulation, so that the storage system is able to handle the next periods.

In order to satisfy a minimum production to operate the data center, a new constraint is added to the linear program as defined in Equation (4.37):

$$Pprod_k \geqslant Pprodmin \tag{4.37}$$

Thus, the power produced in each time step k by the different sources should be greater or equal to the minimal production.

4.3.2.3/ Obtained model

The linear program is able to produce a maximum non constant profile using the storage system composed of batteries and hydrogen system is then written as in the Equation 4.38

4.3.2.4/ Results

Results presented in Figure 4.6 are obtained by applying this linear program. The production $Pprod_k$ is variable in some time step k and greater than the minimal production $Pprodmin$ (see for example at t = 60 h). The same power architecture is used as in the previous subsections. The minimal production is equal to $1MW$ in order to make the variations of $Pprod$ more noticeable. The batteries assure the smoothing as in the last linear program and the level of the battery comes back to the intended one. The level of hydrogen at the end of the horizon is greater than the hydrogen level target fixed, then, the constraint (4.33) is respected.

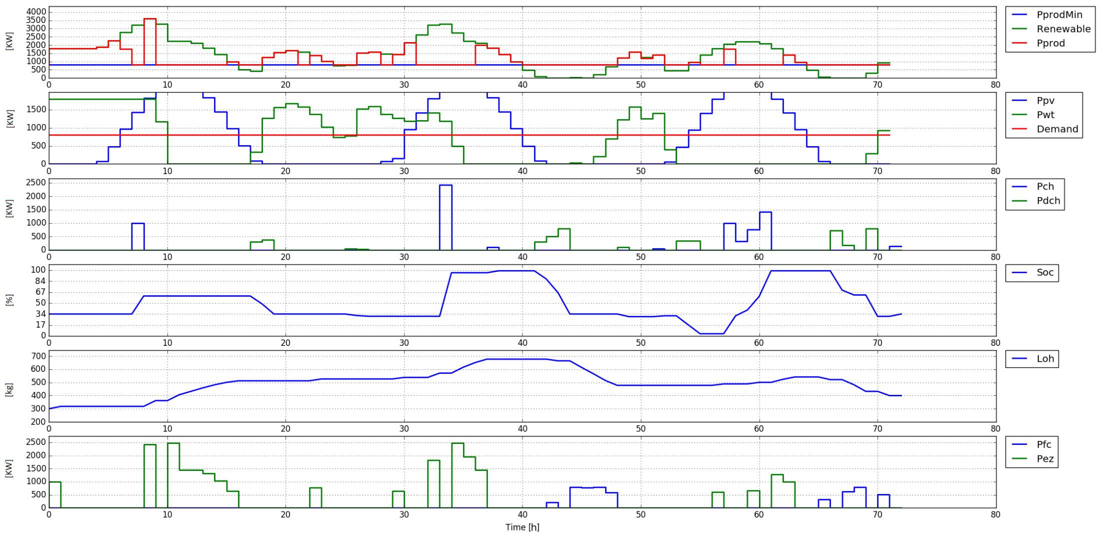


Figure 4.6: Maximum Power Profile management

$$
\left\{
\begin{array}{ll}
\begin{array}{rcl}
 & & \text{maximize} \displaystyle\sum_{k=0}^{K-1} Pprod_k \\
s.t.: & & \\
Pprod_k & \leqslant & Pwt_k + Ppv_k + (Pfc_k + Pdch_k)\eta_{inv} - (Pez_k + Pch_k)\eta_{inv} \\
SOC_k & = & SOC_{k-1}(1-\sigma) + \frac{\eta_{ch} Pch_{k-1}\Delta t - Pdch_{k-1}\Delta t/\eta_{dch}}{Cbat} \\
LOH_K & \geqslant & LOHtarget_D \\
Pprod_k & \geqslant & Pprodmin \\
Pez_k & = & HHVh_2 \times Qez_k/\eta_{ez}/\Delta t \\
Pfc_k & = & LHVh_2 \times Qfc_k \times \eta_{fc}/\Delta t \\
LOH_k & = & LOH_{k-1} + Qez_{k-1} - Qfc_{k-1}/\eta_{tank} \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \geqslant & 0 \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \leqslant & Pch'_k \\
Pch_k & \geqslant & Pch'_k - (1 - x_k)Pchmax \\
Pdch_k & \leqslant & (1 - x_k) \times Pdchmax \\
Pdch_k & \geqslant & 0 \\
Pdch_k & \leqslant & (1 - x_k)Pdchmax \\
Pdch_k & \leqslant & Pdch'_k \\
Pdch_k & \geqslant & Pdch'_k - x_k Pdchmax \\
Pez_k & \leqslant & Pez'_k \\
Pez_k & \geqslant & 0 \\
Pez_k & \leqslant & y_k \times Pezmax \\
Pez_k & \geqslant & Pez'_k - (1 - y_k)Pezmax \\
0 & \leqslant & Pez'_k \leqslant Pezmax \\
Pez_k & \geqslant & y_k \times Pezmin \\
Qez_k & \leqslant & Qez'_k \\
Qez_k & \geqslant & 0 \\
Qez_k & \leqslant & z_k \times Qezmax \\
Qez_k & \geqslant & Qez'_k - (1 - z_k)Qezmax \\
0 & \leqslant & Qez'_k \leqslant Qezmax
\end{array}
&
\left\{
\begin{array}{ll}
Qfc_k & \leqslant Qfc'_k \\
Qfc_k & \geqslant 0 \\
Qfc_k & \leqslant (1 - z_k) \times Qfcmax \\
Qfc_k & \geqslant Qfc'_k - z_k \times Qfcmax \\
0 & \leqslant Qfc'_k \leqslant Qfcmax \\
u_k & \leqslant x_k \\
u_k & \leqslant y_k \\
0 & \leqslant 1 - x_k - y_k + u \\
u_k & \geqslant 0 \\
v_k & \leqslant x_k \\
v_k & \leqslant z_k \\
0 & \leqslant 1 - x_k - z_k + v \\
v_k & \geqslant 0 \\
Bounds: & \\
\forall k \in [\![0, K-1]\!] & Pfc_k \leqslant Pfcmax \\
\forall k \in [\![0, K-1]\!] & Pez_k \geqslant Pezmin \\
\forall k \in [\![0, K]\!] & SOCmin \leqslant SOC_k \leqslant SOCmax \\
\forall k \in [\![0, K]\!] & 0 \leqslant LOH_k \leqslant LOH_{max} \\
\forall k \in [\![0, K-1]\!] & Pch_k, Pdch_k, Pez_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Pch'_k, Pdch'_k, Pez'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Qez_k, Qfc_k, Qez'_k, Qfc'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & x_k, y_k, z_k \in \{0, 1\}
\end{array}
\right.
\end{array}
\right.
\tag{4.38}
$$

4.3.3/ Matching with a requested profile

In this subsection, the aim is to obtain a profile which is matching with a profile requested by the negotiation module to end up with the negotiation between both modules. Thus, the PDM takes as an input the data of the profile requested by the negotiation module, the meteorological data to determine the production, the storage situations, etc.

As an output, the PDM must send more or less a similar profile to the one received by the negotiation module.

To make this possible, some modifications are set for the basic linear program obtained and a new objective function and new constraints are added. To simplify it, the new variable added is explained in Table 4.3.

Table 4.3: Notations

Variable	**Description**
rf	Relaxation factor
λ	Constant value

4.3.3.1/ Objective function

The objective function consists in maximizing the power production $Pprod_k$ hour by hour to increase the chance to fulfill the data center power demand $Pload_k$. Moreover, the long term storage LOH_K should be strongly maintained in order to avoid power waste if $Pprod_k$ is too high (bigger than the load). The difference should be stored in hydrogen. Thus, a high value for λ is necessary to offset the fact that some of the energy produced is lost due to electrolyser and fuel cell efficiency. The problem is thus mathematically translated to:

$$\text{maximize} \sum_{k=0}^{K-1} Pprod_k + \lambda LOH_K \tag{4.39}$$

4.3.3.2/ Added constraints

As in the two previous cases, considering common sense usage rules turns Equation (4.1) into Equation (4.36). The constraint (4.33) is added in order to maximize the hydrogen level at the end of the period of simulation so the storage system is able to handle the next period. In this stage of the negotiation, in order to satisfy the profile requested by NM, a new constraint is added to the linear program as defined in Equation (4.40). Therefore, we introduce a rf provided by the negotiation module and enable the power produced $Pprod_k$ in each time step k to be inside an interval $[(1-rf) \times Pload_k, Pprod_k]$.

$$Pprod_k \geqslant (1-rf) \times Pload_k \tag{4.40}$$

In case of no existing solutions due to power production deficiency to supply the data center, the relax factor value between $0 \leqslant rf \leqslant 1$ is computed using a binary search approach in order to degrade the power response for the data center as less as possible. As the data center load is relaxed, another management is proposed. The solution is thus sub-optimal compared to the initial power demand, but an alternative feasible solution is proposed to overcome the intermittent nature of renewable sources. The evaluation of this degradation impact is evaluated in the following chapter.

The feasibility of the linear program depends on the rf value:

- if $rf = 0$ it means that the profile generated must be the same as the one provided by the NM
- if $rf = 1$ it means that the profile could be entirely different from the one provided

4.3.3.3/ Obtained model

The linear program able to produce a profile that matches with a given one from the negotiation module using the storage system is then written as mentioned in the equation 4.41.

$$
\left\{
\begin{array}{rcl}
\text{maximize} & & \sum_{k=0}^{K-1} Pprod_k + \lambda LOH_K \\
s.t.: & & \\
Pprod_k & \leqslant & Pre_k \eta_{inv} + (Pfc_k + Pdch_k)\eta_{inv} - (Pez_k + Pch_k)\eta_{inv} \\
Pre_k & = & Pwt_k + Ppv_k \\
SOC_k & = & SOC_{k-1}(1-\sigma) + \frac{\eta_{ch} Pch_{k-1}\Delta t - Pdch_{k-1}\Delta t/\eta_{dch}}{Cbat} \\
Pprod_k & \geqslant & (1-rf) \times Pload_k \\
Pez_k & = & HHVh_2 \times Qez_k/\eta_{ez}/\Delta t \\
Pfc_k & = & LHVh_2 \times Qfc_k \times \eta_{fc}/\Delta t \\
LOH_k & = & LOH_{k-1} + Qez_{k-1} - Qfc_{k-1}/\eta_{tank} \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \geqslant & 0 \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \leqslant & Pch'_k \\
Pch_k & \geqslant & Pch'_k - (1-x_k)Pchmax \\
Pdch_k & \leqslant & (1-x_k) \times Pchmax \\
Pdch_k & \geqslant & 0 \\
Pdch_k & \leqslant & (1-x_k)Pdchmax \\
Pdch_k & \leqslant & Pdch'_k \\
Pdch_k & \geqslant & Pdch'_k - x_k Pdchmax \\
Pez_k & \leqslant & Pez'_k \\
Pez_k & \geqslant & 0 \\
Pez_k & \leqslant & y_k \times Pezmax \\
Pez_k & \geqslant & Pez'_k - (1-y_k)Pezmax \\
0 & \leqslant & Pez'_k \leqslant Pezmax \\
Pez_k & \geqslant & y_k \times Pezmin \\
Qez_k & \leqslant & Qez'_k \\
Qez_k & \geqslant & 0 \\
Qez_k & \leqslant & z_k \times Qezmax \\
Qez_k & \geqslant & Qez'_k - (1-z_k)Qezmax \\
0 & \leqslant & Qez'_k \leqslant Qezmax \\
Qfc_k & \leqslant & Qfc'_k \\
Qfc_k & \geqslant & 0 \\
Qfc_k & \leqslant & (1-z_k) \times Qfcmax \\
Qfc_k & \geqslant & Qfc'_k - z_k \times Qfcmax \\
0 & \leqslant & Qfc'_k \leqslant Qfcmax \\
u_k & \leqslant & x_k \\
u_k & \leqslant & y_k \\
0 & \leqslant & 1 - x_k - y_k + u \\
u_k & \geqslant & 0 \\
v_k & \leqslant & x_k \\
v_k & \leqslant & z_k \\
0 & \leqslant & 1 - x_k - z_k + v \\
v_k & \geqslant & 0
\end{array}
\right.
\quad
\left\{
\begin{array}{ll}
Bounds: & \\
\forall k \in [\![0, K-1]\!] & Pfc_k \leqslant Pfcmax \\
\forall k \in [\![0, K-1]\!] & Pez_k \geqslant Pezmin \\
\forall k \in [\![0, K]\!] & SOCmin \leqslant SOC_k \leqslant SOCmax \\
\forall k \in [\![0, K]\!] & 0 \leqslant LOH_k \leqslant LOH_{max} \\
\forall k \in [\![0, K-1]\!] & Pch_k, Pdch_k, Pez_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Pch'_k, Pdch'_k, Pez'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Qez_k, Qfc_k, Qez'_k, Qfc'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & x_k, y_k, z_k \in \{0, 1\}
\end{array}
\right.
\tag{4.41}
$$

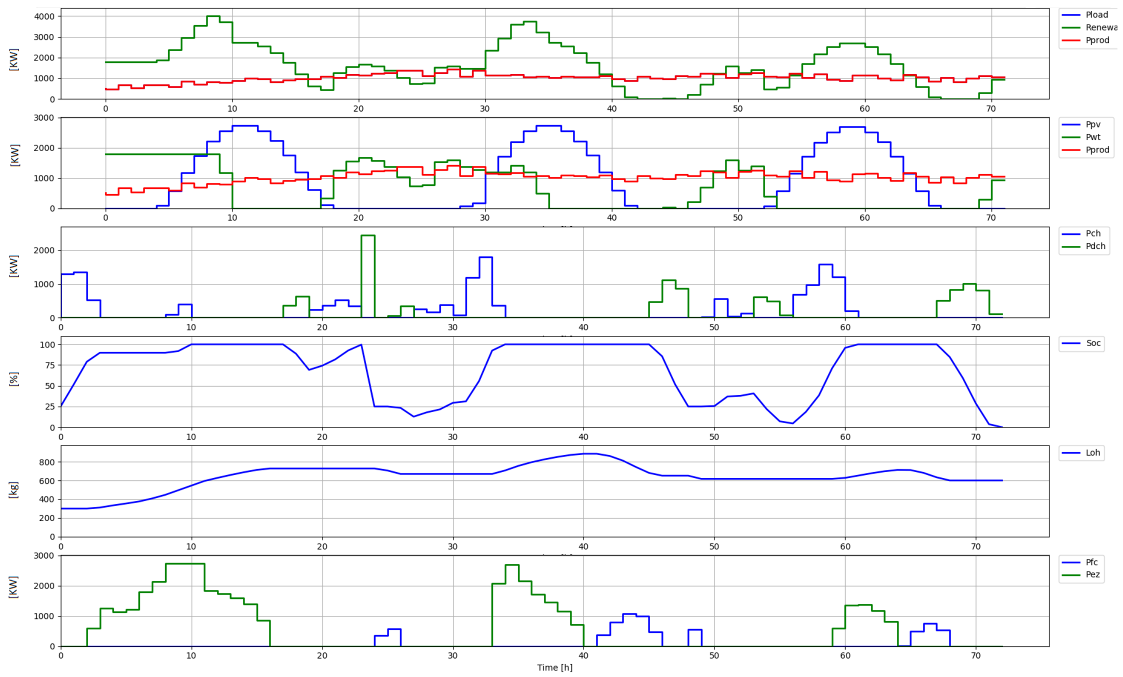


Figure 4.7: Management of the resources to match with a requested profile with $rf = 0$

4.3.3.4/ Results

Results presented in Figure 4.7 are obtained by applying this linear program. The production $Pprod_k$ is completely similar to the demand $Pload_k$ (here the demand represents the power profile received from the negotiation module) as the relax factor is equal to 0. The same power architecture is used as in the previous section. The load is variable between 500 kW and 2 MW in this case in order to see better variation of the profile. The battery assures the smoothing of the day/night alternation and come back to the same level each 24h. At the end of the horizon, starting from the hour 64, one can see an underproduction, the battery and the fuel cell are then working in order to reach the demand. Nevertheless, the constraint (4.33) is maintained and level of hydrogen at the end of the horizon is bigger to the target fixed 300 kg

If one considers a $rf = 0.3$, the results of the MILP are as showed in the figure 4.8. As we can see it, the profile generated is different from the provided one and respects the constraint given in Eq 4.40. If we compare Figures 4.7 and Figure 4.8, the main difference is in the production provided $Pprod$ and in the storage behaviour. In this case, one can see that the storage state of the battery is similar to the previous case. Nevertheless, the hydrogen level is varying and is equal to the target fixed at the end of the horizon $LOH_K = 300\,kg$. This is due to the freedom given by the relax factor.

4.3.4/ Source commitment for the real plateform

In this subsection, the PDM is in charge to find with the best assignment of the sources for the real platform of the profile agreed on between the negotiation, IT and power decision modules. Thus, the PDM takes as an input the data of the profile agreed on from the negotiation module, the meteorological data to determine the production, the storage situations.

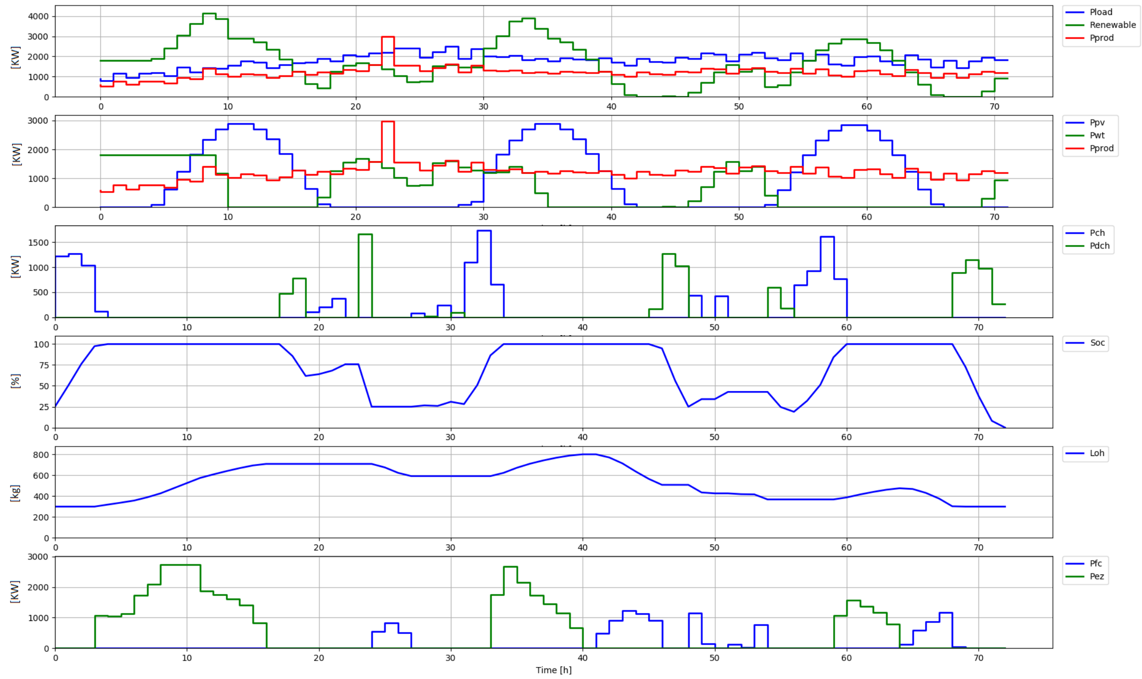


Figure 4.8: Management of the resources to match with a requested profile with $rf = 0.3$

4.3.4.1/ Objective function

The objective function is to maximize long-term stocks by setting a target level hydrogen to reach at each end of the time horizon. Target stocks are set by the previsions of renewable energy production and by the capabilities of the components (Sizing). The problem is thus mathematically translated by minimizing the difference in energy between the hydrogen level calculated at the end of the horizon $\mathcal{H}$ (LOH_K) and a hydrogen target level ($LOHtarget_D$).

The objective function is defined as:

$$\text{minimize} \quad LOHtarget_D - LOH_K \tag{4.42}$$

4.3.4.2/ Obtained model

The linear program obtained for the source commitment is defined, $\forall k \in [\![0, K-1]\!]$ for all constraints except $k \in [\![0, K]\!]$ for constraints containing variables SOC_k, LOH_k, Qez_k, Qfc_k, Qez'_k and Qfc'_k as mentioned in the Equation 4.43.

4.3.4.3/ Results

The figure 4.9 shows the management of the different resources following the constraints explained in the previous sections. The power profile agreed on after the negotiation will be designed with the best assignment of the sources. One can also see that all the constraints are respected, especially for the mutual exclusion between hydrogen and battery. Finally, we notice that the level of hydrogen is decreasing at the end of the horizon $\mathcal{H} = 72$ h as well as the level of the battery because of the lack of renewable production.

$$
\left\{
\begin{array}{rcl}
 & \text{minimize} & LOHtarget_D - LOH_K \\
s.t.: & & \\
Pload_k & \leqslant & Pwt_k + Ppv_k + (Pfc_k + Pdch_k)\eta_{inv} - (Pez_k + Pch_k)\eta_{inv} \\
SOC_k & = & SOC_{k-1}(1-\sigma) + \frac{\eta_{ch} Pch_{k-1}\Delta t - Pdch_{k-1}\Delta t/\eta_{dch}}{Cbat} \\
Pez_k & = & HHVh_2 \times Qez_k/\eta_{ez}/\Delta t \\
Pfc_k & = & LHVh_2 \times Qfc_k \times \eta_{fc}/\Delta t \\
LOH_k & = & LOH_{k-1} + Qez_{k-1} - Qfc_{k-1}/\eta_{tank} \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \geqslant & 0 \\
Pch_k & \leqslant & x_k \times Pchmax \\
Pch_k & \leqslant & Pch'_k \\
Pch_k & \geqslant & Pch'_k - (1-x_k)Pchmax \\
Pdch_k & \leqslant & (1-x_k) \times Pdchmax \\
Pdch_k & \geqslant & 0 \\
Pdch_k & \leqslant & (1-x_k)Pdchmax \\
Pdch_k & \leqslant & Pdch'_k \\
Pdch_k & \geqslant & Pdch'_k - x_k Pdchmax \\
Pez_k & \leqslant & Pez'_k \\
Pez_k & \geqslant & 0 \\
Pez_k & \leqslant & y_k \times Pezmax \\
Pez_k & \geqslant & Pez'_k - (1-y_k)Pezmax \\
0 & \leqslant & Pez'_k \leqslant Pezmax \\
Pez_k & \geqslant & y_k \times Pezmin \\
Qez_k & \leqslant & Qez'_k \\
Qez_k & \geqslant & 0 \\
Qez_k & \leqslant & z_k \times Qezmax \\
Qez_k & \geqslant & Qez'_k - (1-z_k)Qezmax \\
0 & \leqslant & Qez'_k \leqslant Qezmax \\
Qfc_k & \leqslant & Qfc'_k \\
Qfc_k & \geqslant & 0
\end{array}
\right.
\qquad (4.43) \qquad
\left\{
\begin{array}{ll}
Qfc_k \leqslant (1-z_k) \times Qfcmax & \\
Qfc_k \geqslant Qfc'_k - z_k \times Qfcmax & \\
0 \leqslant Qfc'_k \leqslant Qfcmax & \\
u_k \leqslant x_k & \\
u_k \leqslant y_k & \\
0 \leqslant 1 - x_k - y_k + u & \\
u_k \geqslant 0 & \\
v_k \leqslant x_k & \\
v_k \leqslant z_k & \\
0 \leqslant 1 - x_k - z_k + v & \\
v_k \geqslant 0 & \\
Bounds: & \\
\forall k \in [\![0, K-1]\!] & Pfc_k \leqslant Pfcmax \\
\forall k \in [\![0, K-1]\!] & Pez_k \geqslant Pezmin \\
\forall k \in [\![0, K]\!] & SOCmin \leqslant SOC_k \leqslant SOCmax \\
\forall k \in [\![0, K]\!] & 0 \leqslant LOH_k \leqslant LOH_{max} \\
\forall k \in [\![0, K-1]\!] & Pch_k, Pdch_k, Pez_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Pch'_k, Pdch'_k, Pez'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & Qez_k, Qfc_k, Qez'_k, Qfc'_k \geqslant 0 \\
\forall k \in [\![0, K-1]\!] & x_k, y_k, z_k \in \{0, 1\}
\end{array}
\right.
$$

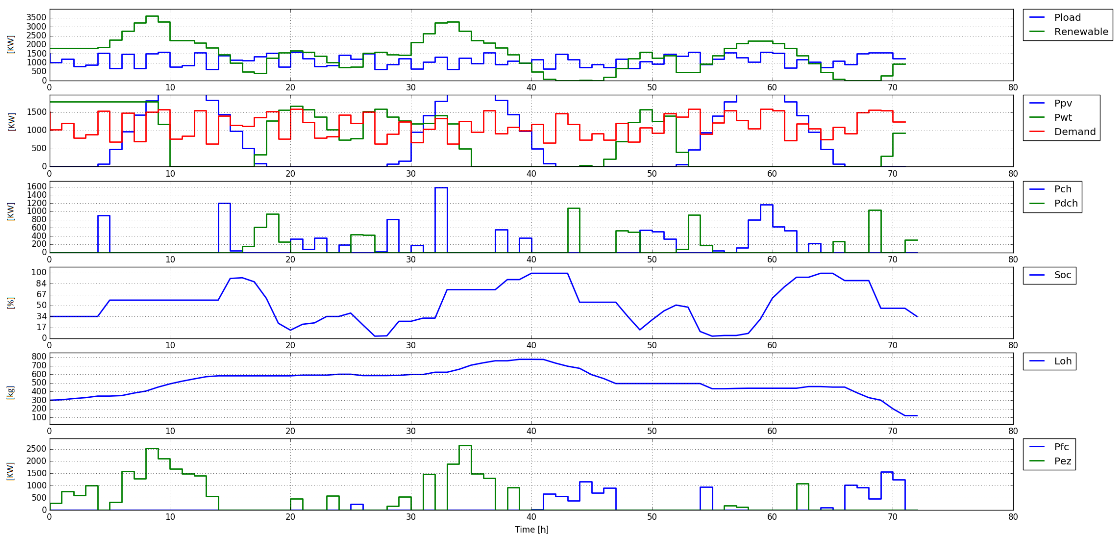


Figure 4.9: Source commitment results

4.3.5/ Summary of the results

The results above suggest several interesting observations:

- The sizing of the electrical sources is very important as well as the values of the initial state of storage. They allow a good running of the programs. Otherwise, the management tool can provide an unfeasible model.
- The mixed integer linear program that matches with a given profile depends extremely of the relax factor provided from the negotiation module. Sometimes, given the storage state at the moment of the negotiation, the management tool provides an unfeasible model and is not capable to answer to the demand. The negotiation takes an action here and sends a second profile with a different relax factor. The increment value of the relax factor is equal to 0.01.
- The analysis of the curves of the MILP "Match Profile" shows an average gap between the curves $Pprod_k$ and $Pload_k$ as shown in following table:

Table 4.4: Gap calculation between $Pprod$ and $Pload$

rf	$\sum Pprod$ [kW]	$\sum Pload$ [kW]	$\overline{gap}$ [kW/h]	Period
0	63820.0	63820.0	0.00	72 h
0.4	65532.0	63820.0	23.78	72 h
0.8	67922.0	63820.0	56.98	72 h
1	68993.4	63820.0	71.85	72 h

- The results of the table 4.4 are obtained after an execution of the MILP during 72 hours with a time step δt equal to 1 hour. The choice of 72 hours is juste an illustrating exemple. It respects the constraint 4.40. This constraint force the program to always be in the interval given.
- Experimentation are performed on an Intel® Core™ i5-6200U CPU @ 2.30GHz × 4 , 8GB RAM, 64-bit using Ubuntu 16.04 LTS as exploitation system. The solving time for each the algorithms is at the scale of seconds which is compatible with the negotiation process as it's a synchronous negotiation.
- The PDM is an efficient management tool of all energy sources (wind turbine, photovoltaics, batteries, fuel cell, electrolyzers) that is able to provide data center demand with 100% renewable energy.

4.4/ Conclusion

In this chapter, we presented (1) an overview of the PDM within the ANR DATAZERO project was made in order to understand the steps of negotiation and all the goals needed to to completely provide a data center with only renewable energy; (2) the usage limits of the different electrical devices (constraints) (3) the linearization of the constraints in order to get a convenient usage of the sources and the mixed linear program which is used in the power decision module (PDM); (5) the power decision module which is a management tool able to manage a hybrid renewable energy system presenting four optimization problems depending on the stage of the negotiation. Additional constraints are also written to allow us to find appropriate, exact and optimal solutions. These problems have been illustrated by several simulations. We show that our optimal models are convenient to address the power commitment of an 100% renewable energy data center, since the resolution is obtained within few seconds. The solving time for each the algorithms is at the scale of seconds even for horizons of one year with a time step of hours.

Chapter 5

Experiments

The last two chapters have explained the sizing tool and management tool used in the electrical infrastructure. This chapter aims at integrating both optimization modules (sizing and PDM) in order to regulate the power energy production and determine how to produce and consume energy within the whole system under the end-user service level constraint. Moreover, an exploitation of these tools will be used in order to show a set of scenarios against which the project will be challenged.

Contents

5.1/ Objectives & motivations

In the precedent chapters, we proved that the sizing module is able to provide an electrical infrastructure based on meteorological data and a power load. This architecture is then used by the power decision module in order to find the an optimized comitment of the sources following the chosen goal.

In this chapter, experiments are done in order to measure the usage scope of both sizing and power decision (PDM) modules on data center lifetime. These experiments will be done year after year in order to find the best assignment of the electrical sources. This optimal management has to meet our expectations in terms of satisfying the power demand while using the renewable energy and limiting the storage losses.

In fact, the targeted data center is intended to be used all year long without resorting to the conventional power grid. In order to make it possible, the latter should maintain an energy level in its long-term storage in order to be able to operate even if days where the power from renewable sources is not enough to address the power demand. Indeed, the idea is that the most favorable days (let say in summer) has to compensate for the worst days (let say in winter). The challenge is therefore to smooth consumption over the year thanks to long-term storage (H_2) and over days thanks to the short term storage (batteries). It is then a question of scheduling a commitment of the sources, and thus managing both the short term and long term storage.

5.2/ Methodology

In the power decision module, four MILP were presented in order to reach the goal fixed. Nevertheless, in this section, the mixed integer linear program (MILP) chosen to be used in the experiments is explained. In a second time, a reminder of the evaluation performance metric is given. An assumption that the weather forecast is reliable is made in the following. Then, the study of the confidence interval related to the weather forecast is not taken into account.

5.2.1/ Mixed integer linear program used

The most used optimization problem in the negotiation loop is the one that tries to match with a requested profile from the negotiation (4.41). It tends to maximize the power production $Pprod_k$ with maintaining the level of hydrogen LOH_K at the end of the horizon $\mathcal{H}$ to make possible the seasonal compensation without wasting energy and to be able to handle the next season.

Moreover, the long term storage LOH_K should be strongly maintained rather than allowing waste if overproduction is too high. Thus, a high value for λ is necessary to offset the fact that some of the energy produced is lost due to electrolyser and fuel cell efficiency. Also, in order to match $Pprod_k$ with the data center load $Pload_k$ for each time step k, a relax factor rf is used as in the constraint (5.1).

$$Pprod_k \geqslant (1 - rf) \times Pload_k \tag{5.1}$$

5.2.2/ Evaluation performance Metrics

In order to evaluate the solutions obtained by the mixed integer linear program (MILP), some new metrics are settled down in this section. Some metric presented in the section 3.4 of the “Sizing Strategy” chapter are also used in order to rate the obtained solutions such as the loss of power supply probability (LPSP) and the level of autonomy (LA).

5.2.2.1/ Unused renewable energy (URE)

As the charging and discharging power of both system batteries and hydrogen are constants, the MILP enables, sometimes, to not use all the energy produced by the renewable sources. Thus, this quantity of unused renewable energy is computed as:

$$URE = \frac{1}{\mathcal{H}} \sum_{k=0}^{K-1} \left(\max(Pprod_k - Pload_k, 0) \times \Delta t \right) \tag{5.2}$$

5.2.2.2/ Percentage of the Energy Produced to demand (PEP)

It consists in the computation of the power produced by the electrical infrastructure $Pprod_k$ divided by the demand. This is mathematically proceeded by the following expression:

$$PEP = \frac{\sum_{k=0}^{K-1} \min(Pprod_k, Pload_k)}{\sum_{k=0}^{K-1} Pload_k} \tag{5.3}$$

5.3/ Experiments and discussions

First, the sizing module takes as inputs the weather conditions, power demand $Pload_k$ and determine the best configuration needed based on these informations. The data representing the solar radiation and wind speed have been measured on an hourly scale from January 2004 till December 2012 in a coastal area. To be more precise, the endogenous data of the solar radiation and wind speed time series have been measured at Los Angeles (Latitude: 34.57, Longitude: -118.02, Time Zone: -8). These data has been obtained respectively from the National Solar Radiation DataBase (NSRDB) [116] and from wind prospector of the National Renewable Energy Laboratory (NREL) [43]. From the IT part, the data center demand were obtained from the traces of access logs from the World Cup web site of 1998 [23]. This load choice is made because of its similarity to the traces of a real data center. Based on these data, a sizing study was made and an average infrastructure as showed in the table 5.1 is given. This sizing is an input the MILP needed to be solved.

Table 5.1: Infrastructure sizing used in simulations

Nwt	Npv	Pch	Pdch	SOCmax	Pez	Pfc	H_2
2	6650	980	555	5287	910	832	6382

Then, the electrical power system configuration like the number of wind turbines Nwt, the solar panel surface Npv, battery storage capacity $SOCmax$, the capacity of the hydrogen tank H_2, the efficiency of each component are sent to the PDM in order to start optimization. As an output, the solver gives a schedule of each time slot for supply-side source and energy storage usage, according to the availability of renewable power.

The same infrastructure sizing will be used in many management scenarios with a different resolution time window for a better illustration of the difference between the obtained results. A resolution time window consists in iteratively solving the power management program for a given period length $\mathcal{H}$ as many times as it is needed to cover the whole year. First, we proceed to the verification of the proper functioning of the optimization strategy.

5.3.1/ Management results

Figure 5.1 shows a solution obtained by solving the MILP on a given period $\mathcal{H} = 14$ days $= 336\,h$ with a time slot $\Delta t = 1\,h$ in the winter during the year 2010. This year was chosen for a better illustration of the results (compared to the other years). This solution consists of the evolution of each variable of the program, hour by hour, each visible on a specific window of Figure 5.1. The first windows shows the data center demand $Pload$ and the produced power $Pprod$, used to supply the data center, optimized by the MILP. The second window shows in details the the solar panels Ppv, the wind turbines Pwt and the total Pre produced. The third window shows both the charging Pch and discharging power $Pdch$

of the batteries followed by the variation of state of charge of the batteries (SOC). The last two windows of Figure 5.1 show the actual level of hydrogen LOH computed by the MILP. Finally, the electrolyzer power Pez and fuel cell power Pfc obtained by the MILP is shown to explain the evolution of LOH. This figure was chosen in winter in the worst

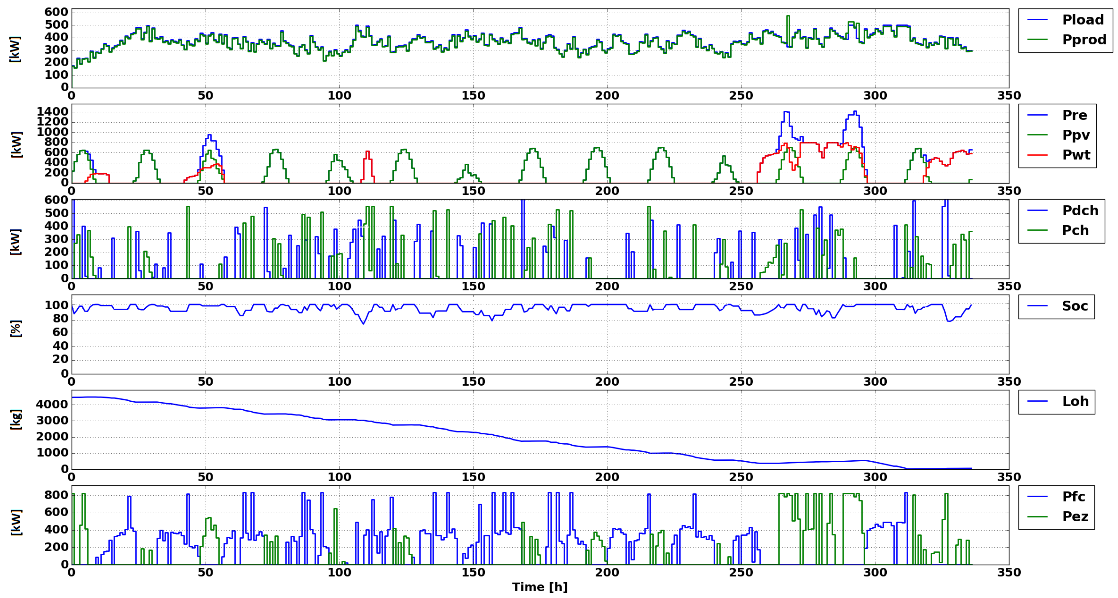


Figure 5.1: Management of the renewable sources for 15 days with $rf = 0.02$

conditions of the year 2010 to show the aim of the relax factor. In fact, with an optimal relax factor $rf = 0.02$, the power supplied to the data center is relaxed (i.e., reduced as less as possible) while a solution is possible because of the problem constraints. The first founded solution is then displayed. It is used in order to propose an alternative power profile if no solution exists to answer to the data center power demand. Thus, the latter could postpone or reschedule the jobs. Based on Figure 5.1, one can see that the battery state of charge comes back to its initial level each 24 hours because batteries are only responsible of short term energy storage (daily compensation) to ensure the hourly lack of renewable energy (e.g., solar night and day alternation) during the same day.

Also, as it is responsible of the seasonal offset, one can see that the level of hydrogen is decreasing as there is not enough renewable energy to ensure the demand (no wind) and get almost empty at the end of window.

Thus, as the linear program is working properly, the following subsection shows different usage of this program. Indeed, the management program should be able to find a solution under any weather conditions. The metrics explained in Section 5.2.2 are used to verify the robustness of our approach.

5.3.2/ Management scenarios results

In this part, the program is executed following different resolution time windows for the same horizon (1 year) under different weather conditions. In fact, the same sizing is played from the year 2004 till 2012. For each year, several resolution time windows are simulated with the weather condition of the year. One simulation consists in iteratively solving the power management program for a given period length $\mathcal{H}$ as many times as it is needed to cover the whole year. This resolution time window $\mathcal{H}$ is related to the weather prediction time.

Different cases of resolution time window are exploited to shows the impact of the decision on the management proposed:

- 1st case: resolution window is equal to the horizon fixed (1 resolution for 1 year) (Figure 5.2)
- 2nd case: resolution window is equal to 1 week (52 resolutions for 1 year) (Figure 5.3)
- 3rd case: resolution window is equal to 3 days (121 resolutions for 1 year) (Figure 5.4)
- 4th case: resolution window is equal to 1 day (365 resolutions for 1 year) (Figure 5.5)

The aim of this comparison is to choose the best management time window using this approach (Should the decision be taken only 1 time per week, per three days or per days?). The chosen metrics are applied on each cases study. For all cases, an average value of the metrics is computed and displayed in the tables. In the following tables, the column loss of power supply probability (LPSP) is divided on two columns, the average and the standard deviation of the LPSP values obtained for each resolution in different simulations.

Experimentation are performed on an Intel® Core™ i5-6200U CPU @ 2.30GHz × 4 , 8GB RAM, 64-bit using Ubuntu 16.04 LTS as exploitation system. The solving time for each year with different resolution time windows will be also compared.

5.3.2.1/ One resolution for one year

Figure 5.2 shows the obtained solution of the solver. The figure shows 5 subfigures. The first presents the power load of the data center Pload and the optimized power produced by the infrastructure Pprod. The second window presents the produced power by the photovoltaic panel Ppv, wind turbine Pwt and their sum to identify the renewable primary power produced Pre. The third and fourth windows are dedicated to the battery where the charging Pch, discharging power Pdch and the state of charge Soc are illustrated. The last two windows are reserved to hydrogen storage system where the the level of hydrogen calculated by the MILP (Loh), the level of hydrogen determined by the sizing study (loh-Target), the fuel cell and electrolyzer power respectively Pfc and Pez are identified.

Based on this figure, one can see that the produced power is not able to answer to the demand from the beginning of the time window till the hour 3500. In this period, the level of hydrogen computed by the solver is really close the loh-Target. From the hour 3500 till 6500 one can see that the production is bigger than the load demand.

The table 5.2 resume the metric results for a one year resolution with $\mathcal{H} = 1$ year $= 8760\,h$. The results of the relax factor shows quite high values as only one decision is taken for the whole year. This high relax factor has an impact on the LPSP which shows that, for example, during 2005, 77% of the data center demand, on the considered horizon, was not supplied by the system. Nevertheless, the infrastructure was able to give 94% of the energy produced to the demand during the year. It means that the difference between the load and the provided production is not big. Also, one can see that the level of autonomy

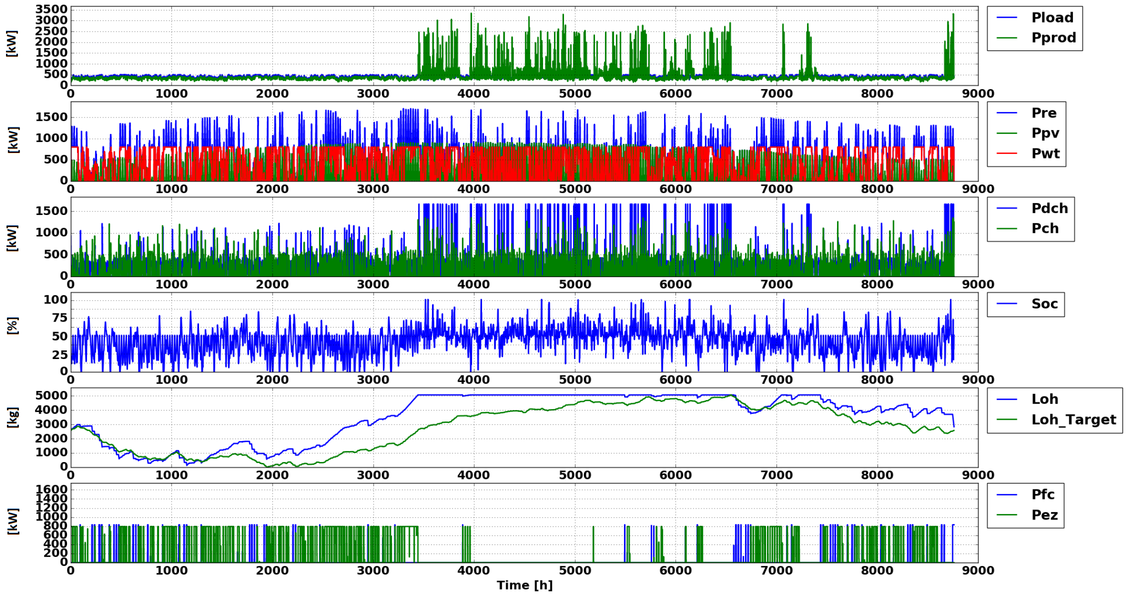


Figure 5.2: Management of the renewable sources in the 1st case for the year 2008

Table 5.2: Metric for a one year simulation with $\mathcal{H} = 1$ year $= 8760\,h$

year	relaxation	LA	LPSP		URE	PEP
			Average	SD		
2004	0.07	0.67	0.98	0	3.45	0.93
2005	0.08	0.69	0.77	0	84.94	0.94
2006	0.11	0.66	0.85	0	47.08	0.90
2007	0.31	0.64	0.75	0	82.54	0.77
2008	0.18	0.64	0.85	0	43.84	0.78
2009	0.27	0.63	0.83	0	62.16	0.78
2010	0.48	0.60	0.72	0	121.51	0.67
2011	0.43	0.64	0.66	0	155.09	0.73
2012	0.3	0.61	0.79	0	60.94	0.76

LA is around 60% for all the years and that the unused energy is quite different from a year to another.

After analyzing the results, taking only one decision for a whole year is actually not accurate due to the uncertainty of the weather conditions . Then, we decided to shorten the resolution window to 1 week (52 resolutions for one year), 3 days (121 resolutions for one year) and to 1 day (365 resolutions for one year). The obtained solutions are illustrated in the figures (5.3), (5.4), (5.5). The evaluating metric are summarized respectively in the tables 5.3,5.4,5.5

5.3.2.2/ 52 resolutions for one year

In this subsection, the resolution time window is equal to 1 week. It means that during the year, the solver tries to find an optimized power source commitment for each week during the whole year.

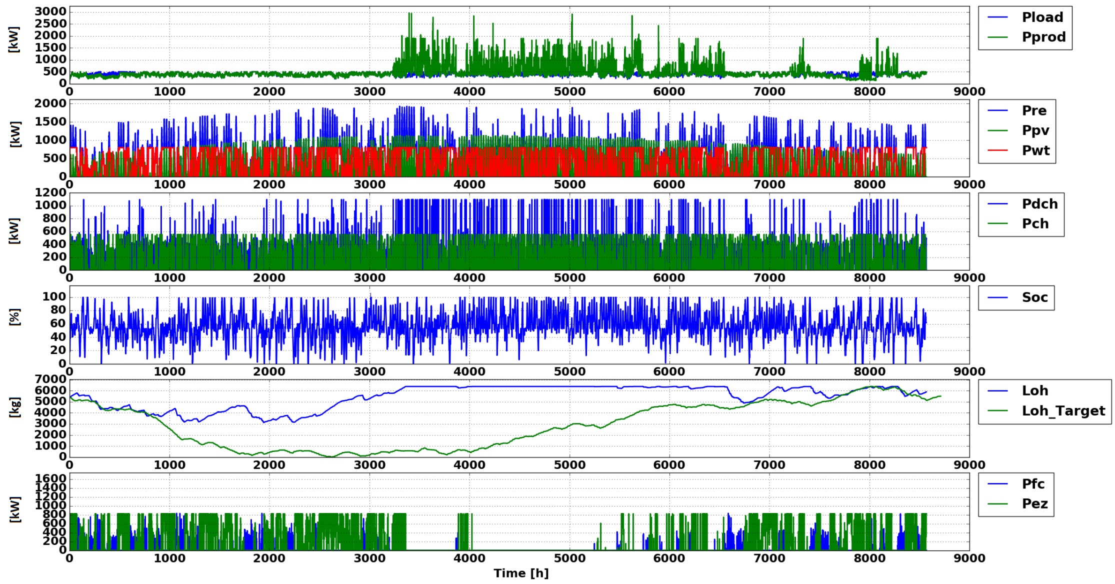


Figure 5.3: Management of the renewable sources in the 2nd case for the year 2008

The solution obtained is displayed on the figure 5.3. One can see that during the winter period (till the hour 3500), the solution is better than in the first case. The energy produced is able to supply the demand over the year except for 2 or 3 weeks (hours around 500 and 8000).

From the other hand, the level of hydrogen computed follows the trends of the loh-target of the sizing but it is not really fitting to the values and is even higher (good year).

Table 5.3: Metric for a one year simulation with $\mathcal{H} = 1$ week $= 168\,h$

year	relaxation	LA	LPSP		URE	PEP
			Average	SD		
2004	0	1	0.06	0.04	84.36	0.67
2005	0	1	0.05	0.04	158.81	0.69
2006	0	1	0.05	0.03	129.46	0.67
2007	0.06	0.93	0.20	0.31	56.70	0.64
2008	0.02	0.97	0.18	0.29	56.89	0.64
2009	0.04	0.95	0.26	0.35	43.21	0.63
2010	0.10	0.89	0.31	0.39	45.65	0.60
2011	0.08	0.91	0.26	0.37	85.30	0.64
2012	0.12	0.93	0.30	0.37	1.55	0.61

Based on the values of the table 5.3, the relax factor is smaller than the first case especially for the years 2004, 2005 and 2006 where the program is able to supply the power demand of the data center without relaxing any week. The loss of power supply is almost equal to 0 for all the years which confirm our hypothesis.

From an opposite site, the percentage of energy produced to demand PEP decreases compares to the first case.

5.3.2.3/ 121 resolutions for one year

In this subsection, the resolution time window is equal to three days. It means that during the year, the solver tries to find an optimized power source commitment for three days during the whole year equal to 121 resolutions for one year. Same ascertainments can

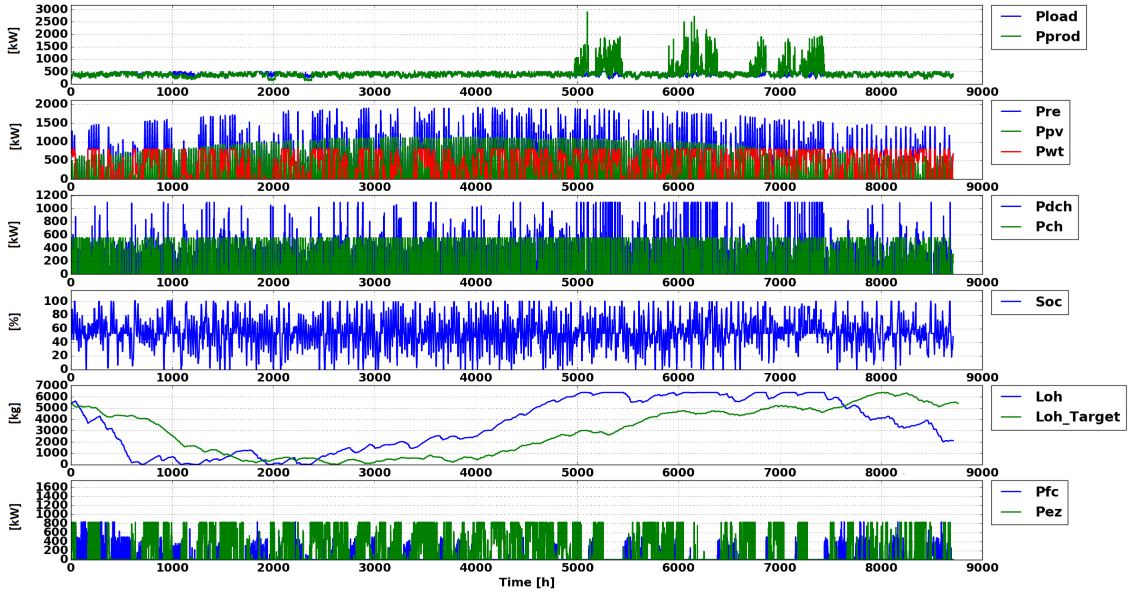


Figure 5.4: Management of the renewable sources in the 3rd case for the year 2008

also be verified here. The solution obtained is displayed on the figure 5.4. One can see that during the winter period (till the hour 3500), the solution is better than the first and second case. The energy produced is able to supply the demand over the year except for 3 days (around the hours 1000 and 2000).

From the other hand, the level of hydrogen computed follows much better the trends of the loh-target determined by the sizing. Moreover, it tries to get the closest possible and to fit to the target values.

Table 5.4: Metric for a one year simulation with $\mathcal{H} = 3$ days $= 72\,h$

year	relaxation	LA	LPSP		URE	PEP
			Average	SD		
2004	0	1	0.06	0.05	84.59	0.67
2005	0	1	0.05	0.05	159.24	0.69
2006	0	1	0.05	0.05	129.81	0.67
2007	0.04	0.95	0.17	0.27	54.05	0.64
2008	0.02	0.98	0.15	0.24	57.52	0.64
2009	0.04	0.95	0.20	0.30	42.42	0.63
2010	0.08	0.91	0.28	0.37	44.27	0.59
2011	0.06	0.93	0.22	0.33	85.46	0.64
2012	0.05	0.94	0.23	0.32	1.55	0.61

Based on the values of the table 5.4, the relax factor is smaller than the first and second cases. In fact, for the years 2004, 2005 and 2006, the architecture is able to supply the power demand of the data center without relaxing any days. For the other years, the relaxation factor is lower than the precedent cases. The loss of power supply is almost the same as the second case ,equal to 0 for all the years.

From an opposite site, the percentage of energy produced to demand PEP decreases compares to the first case.

5.3.2.4/ 365 resolutions for one year

In this subsection, the resolution time window is equal to one day. It means that during the year, the solver tries to find an optimized power source commitment for one day during the whole year equal to 365 resolutions for one year.

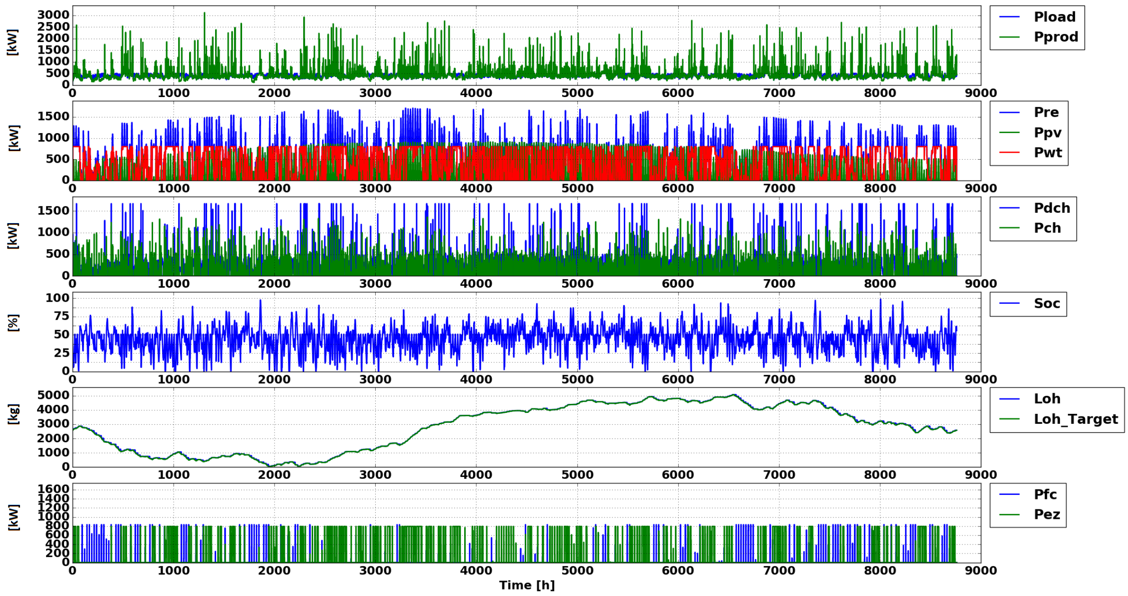


Figure 5.5: Management of the renewable sources in the 4th case for the year 2008

Same ascertainments can also be verified here. The solution obtained is displayed on the figure 5.4. One can see that during the winter period (till the hour 3500), the solution is better than the first and second case. The energy produced is able to supply the demand over the year except for some days (5 days).

From the other hand, the level of hydrogen computed follows much better the trends of the loh-target determined by the sizing. Moreover, it really fits to the target values.

Based on the values of the table 5.5, the relax factor is quite similar to the third case. In fact, for the years 2004, 2005 and 2006, the architecture is able to supply the power demand of the data center without relaxing any days. For the other years, the relaxation factor is lower than the precedent cases. The loss of power supply is almost the same as the third case, equal to 0 for all the years.

From an opposite site, the percentage of energy produced to demand PEP decreases compares to the first case.

Table 5.5: Metric for a one year simulation with $\mathcal{H} = 1$ day $= 24\,h$

year	relaxation	LA	LPSP		URE	PEP
			Average	SD		
2004	0	1	0.07	0.09	84.13	0.67
2005	0	1	0.05	0.06	158.37	0.69
2006	0	1	0.06	0.07	129.09	0.66
2007	0.03	0.96	0.09	0.26	53.61	0.64
2008	0.01	0.98	0.08	0.20	56.71	0.64
2009	0.02	0.97	0.09	0.25	40.72	0.63
2010	0.06	0.93	0.08	0.34	43.98	0.60
2011	0.05	0.94	0.08	0.29	84.65	0.64
2012	0.03	0.96	0.1	0.28	1.55	0.61

5.3.3/ Summary of the results

In the three tables, one can see that the smaller is the resolution window, the smaller is the relax factor. In fact, one can see that for the years 2004, 2005, 2006, rf decreased from around 0.08 to 0 which improved the loss of power supply and made it capable to answer to the data center demand more often. Contrarily, one can see that the percentage of the energy produced by the electrical infrastructure to demand PEP degrades with the smaller time windows. In fact, this is due to the inability of the system to answer to the demand in some time steps (i.e., instead of treating the lack of overproduction during a year, it is treated in a day). Indeed, the LPSP values computed for the different horizon $\mathcal{H} \in \{168, 72, 24\}$ are quite similar. Nevertheless, the standard deviation shows that the smaller is the resolution time window, the bigger is the periodic variation. To conclude, following to the metric results, the constraints of IT scheduling and the uncertainties of the weather prediction, the best resolution time windows is equal to 3 days as showed in Figure 5.4.

Based on Figure 5.4, one can see that around the hour 1000, the system is unable to meet the data center demand $Pload$: the fuel cell started working and the level of hydrogen decreases till reaching 0 kg. In this moment, the $Pload$ is relaxed by rf for those three days. This relaxation could also be translated by a purchase of hydrogen to fulfill the demand. Starting from the hour 4000 (beginning of june), one can see that the level of hydrogen increases till reaching it maximum level at around the hour 5500. $Pprod$ is then bigger than the demand. This unused renewable energy could be translated by selling hydrogen.

The resolution time of the MILP is in the order of seconds. For the one year time windows, the resolution time is equal to $32.48\,s$. The weekly resolutions takes around $40.5\,s$. The three days resolution time is equal to $44.8\,s$ and the daily resolution time is equal to $55.7\,s$.

5.4/ Synthesis

In this chapter, the sizing and power decision modules were exploited under four optimization scenarios depending on the resolution time window. These problems have been illustrated by several simulations. The optimal model convenient to address the power commitment of a 100% renewable energy data center, since the resolution is obtained within few seconds. Nevertheless, these results depends strongly of the meteorological data downloaded. Thus, in order to get better results, a study of a forecasting method is presented in the following chapter.

Conclusions and Future Works

General synthesis

In the context of the increasing electrical consumption of the data center (factories of the digital), using hybrid renewable energy system is one of the most important solution. In fact, it constitutes a strategic asset to win the trade between the energy dependency and the environmental restrictions. The research works proposed in this manuscript has focused on, as part of the ANR DATAZERO project, the sizing and optimal management of hybrid renewable energy for a 100 % green supply of a data center.

First, up-to-date information on electrical consumption and CO2 emission was made on a world wild level. Moreover, a review on renewable energy technologies and energy storage systems technologies were presented in the first chapter. These technologies were compared depending on their maturity, cyclability, energy density, and their costs. The main sources used in our infrastructure are: solar, wind, hydrogen and batteries.

In a second time, an overview on the energy efficiency strategy for green data center was presented, in chapter 2, in order to get positioned in the existing works. Then, a study on data center and the electrical chosen sources model was presented. In fact, two different models were presented for the electrical sources used for two different purposes. Model (A) generally used in the literature for sizing and long term management of renewable energy (hourly scale, daily, scale, etc.). A second model (B) were proposed for the control of the devices and the short term management (second or minute scale). Our choice was to focus on the model (A) following the our fixed goals (Sizing and management of an infrastructure for data center application). The chapter 2 ends up by a review on different strategies adopted by researchers in designing or managing these systems in order to supply power demand.

After consulting the state of art of the published research related to hybrid renewable energy sources, their sizing and management, we proceeded in the chapter 3 to sizing. In fact, in the beginning of this chapter, we presented the principle of our sizing strategy where the problem was taken backward: based on a given infrastructure, we search for the power load that the latter is able to provide over the year. This was done such as the overproduced energy is equal to under produced one taking into account the energy efficiency of the storage. In the second part of this chapter, the final sizing strategy capable to respond to the data center demand is explained. The tool provided different results for each year with several configurations; Full PV, Hybrid 1 composed of 1 wind turbine and the rest of PV panels, Hybrid 2, etc, or Full WT. After introducing and analyzing some metrics such as the annual cost, the renewable energy penetration and the reliability, the best configuration chosen among all is Hybrid 2 as it is the best in quality/services. Once the configuration is fixed, we needed an exact sizing for our data center demand, thus we played different scenarios with three different sizing over the years: best year sizing, worst year and average year sizing in order to verify the storage variation over the years. Thus, the sizing tool was able to deliver the best architecture taking into account the played scenarios, storage variation and while respecting the applied performance criteria was made.

Taking into account the sizing results, we started working on an optimal management of the hybrid renewable energy system, in the chapter 4, in order to supply the data center. An overview of the power decision module, within the ANR DATAZERO project, was made in order to understand the steps of negotiation and all the goals needed. In a second time, the constraints deduced from the models and the usage limits of the different electrical devices were defined. As these constraint were not linear, we proceeded to their linearization to construct a mixed integer linear program (MILP) which is used in the power decision module (PDM). Following the negotiation step, four optimization problems able to assign the sources in order to supply the data center load are made. The simulation results illustrated, have shown a good functioning of the four obtained MILP. To end up, we created a management tool able to manage a hybrid system composed of photovoltaic panels, wind turbines, hydrogen and battery energy systems capable to find 4 different managements following the desired goal.

In the 5th chapter, the sizing and power decision modules were exploited under four optimization scenarios depending on the resolution time window. A resolution time window consists in iteratively solving the power management program for a given period length $\mathcal{H}$ as many times as it is needed to cover the whole year. These problems have been illustrated by several simulations with different time windows (year, week, 3 days and 1 day). Thus, we showed that our optimal models are convenient to address the power commitment of an 100 % renewable energy data center, since the resolution is obtained within few seconds. The solving time for each the algorithms is at the scale of seconds even for horizons of one year with a time step of hours.

As the obtained solutions are quite related to the meteorological data downloaded: where the sizing results were fixed based on the previous meteorological data (year 2004 till 2012). Moreover, the power decision module makes management decision of the infrastructure for 1 year that is challenged every 3 days. Consequently, valid forecasting of meteorological data needs to be studied in order to get the closest assignment to the real data. Chapter 6 presents a comparison among four distinct solar radiation and wind speed generation forecasting models. It is shown that in general, SARIMA model is quite good in the forecasting of the solar radiation during years and fits very well the data because of their seasonal distribution. We also pointed out that the performance of the used model in forecasting the solar during the years is more precise than the ones for the wind speed which degrades noticeably for long term previsions. It is hence important to predict wind speed variation as precisely as possible. This shows the interest to consider other models or characteristics such as Markov Switching ARMA [44] to improve the precision of the results in order to get an optimal sizing and management for the hybrid renewable energy system supplying a data center power demand.

Future works and perspectives

Different perspectives can be considered as a result of this work. First of all, Direct perspectives include additions to improve the results obtained and complete the theoretical study. More general perspectives in a broader context are then proposed.

The sizing study can be improved if some modification on the annual system cost is made. In fact, the cost estimation studies can be ameliorated if we take into account the cable costs (among others) and into consideration an interval of cost variation over the years (inflation).

In the power decision module, aging coefficient of the electrical sources used in the infrastructure were not taken into account. Nevertheless, it is quite important to include the aging in these quid of work as it can completely change the planned management.

Improve both tools by working on forecasting and refining the trends of solar radiation and wind speed using other models. This will minimize the event (power difference) between the planned assignment of the electrical sources and the production of the real infrastructure.

Finally, it would be interesting to take over the different approaches related to the optimization of sizing and management:

- by adding studies of an efficient weather forecasting method. In fact, the obtained results depend strongly of the meteorological data downloaded. valid forecasting of meteorological data need to be studied in order to get the different trends of wind speed and solar radiation during the next years. Based on these trends, the overproduction period and underproduction period can be predefined in order to ensure a sustainable, optimal and green management during the lifetime of the system. In Appendix B, an initiating work is done where one forecasting model was defined and used in order to evaluate the robustness of our approach.
- By taking into account uncertainty directly in the algorithms. As a result of these algorithms, it would provide either a certain level of uncertainty is associated
- By offering robust solutions, which are not the best solutions and not the worst neither. It consists in solutions that are able to provide a good level whatever the scenario.

List of Publications

All the work and results presented in this these have been published in national and international peer-reviewed conferences and in newspapers. Extended versions of some articles are available in research reports.

I have also contributed in writing the deliverables of the ANR DATAZERO project and assisted to the planary meetings which helped me aquire a certain writing disciplines and an analysis vision to the obtained results

Preliminary results were also presented during workshops if the theme of the scheduling entitled "New challenges in Scheduling Theory" was held in Aussois in April 2018.

Journal Articles

[J1] Jean-Marc Pierson, Gwilerm Baudic and Stéphane Caux, Berk Celik, Georges Da Costa, Léo Grange, Marwa Haddad, Jérôme Lecuivre, Jean-Marc Nicod, laurent Philippe, Veronika Rehn-Sonigo, Robin Roche, Gustavo Rostirolla, Amal Sayah, Patricia Stolf, Thi Minh-Thuyen and Christophe Varnier, DATAZERO: Datacenter With Zero Emission and Robust Management Using Renewable Energy, in IEEE Access journal, Volume 7, pages 103209 – 103230, july 2019, https://doi.org/10.1109/ACCESS.2019.2930368

This paper aims at presenting an overview of the ANR DATAZERO project. My contribution in this papier is dedicated to the electrical infrastructure: Electrical Sizing, Optimal management of the hybrid renewable energy.

[J2] Minh-Thuyen Thi, Jean-Marc Pierson, Georges Da Costa, Patricia Stolf, Marwa Haddad, Jean-Marc Nicod, Gustavo Rostirolla, Negotiation Game for Joint IT and Energy Management in Green Datacenters in Future Generation Computer Systems, Elsevier. 2019. (In press)

This paper is under major revision. It aims at presenting an the negotiation process in the ANR DATAZERO project. My contribution in this papier is dedicated to the electrical infrastructure: Optimal management of the hybrid renewable energy.

Current submission:

[J3] Maroua Haddad, Jean-Marc Nicod, Marie-Cécile Péra and Christophe Varnier, Stand-Alone Renewable Power System Scheduling for a Green Data-Center using Integer Linear Programming, in Journal of Scheduling, Springer.

This paper aims at presenting the power decision module of the ANR DATAZERO project. My contribution is dedicated to the Optimal management of the hybrid renewable energy following the chosen goal.

[J4] Gustavo Rostirollan, Haddad Marwa, Grange Léo, Minh-Thuyen Thi, Ayham Kassab, Pierson Jean-Marc, Baudic Gwilherm, Caux Stéphane, Celik Berk, Nicod Jean-Marc, Patricia Stolf, Da Costa Georges, Lecuivre Jérôme, laurent Philippe, Veronika Rehn-Sonigo, Robin Roche, Sizing and Management of Energy Sources for Green Datacenters with Renewable Energy, in Renewable & Sustainable Energy Reviews, Elsevier.

This paper is a survey on Sizing and Management of Energy Sources for Green Datacenters with Renewable Energy. My contribution is dedicated to electrical part of this article: Electrical sources model, Sizing and Optimal management of the hybrid renewable energy.

International conferences

[IC1] Haddad Marwa, Nicod Jean-Marc, Pera Marie-Cecile Hydrogen infrastructure: data-center supply-refueling station synergy in 2017 IEEE Vehicle Power and Propulsion Conference (VPPC), pages 1 –6, Belfort, 2017

This paper aims at presenting a Sizing idea by combining data center and refueling hydrogen station.

[IC2] Haddad Marwa, Nicod Jean-Marc, Varnier Christophe, Péra Marie-Cécile Mixed Integer Linear Programming Approach to Optimize the Hybrid Renewable Energy System Management for supplying a Stand-Alone Data Center, THE Tenthth International GREEN and Sustainable Computing CONFERENCE (IGSC), Alexandria, VA, USA, Co-sponsored by IEEE Computer Society & STC Sustainable Computing, October 2019.

This paper aims at presenting different management scenarios of management of a hybrid system using MILP approach.

[IC3] Haddad Marwa, Nicod Jean-Marc, Boubacar Mainassara Yacouba, Rabehasaina Landy, Al Masry Zeina Péra, Marie-Cécile Wind and Solar Forecasting for Renewable Energy System using SARIMA-based Model, 6th International conference on Time Series and Forecasting (ITISE 2019), Granada, Spain, Contribution to Statistics (Springer), September 2019.

This paper aims at presenting a forecasting method of wind speed and solar radiation using the SARIMA model.

National conferences

[NC1] Haddad Marwa, Nicod Jean-Marc, Christophe Varnier, Pera Marie-Cecile Optimisation de l'engagement des sources pour un Data Center approvisionné en énergie renouvelable Compas'2018 : Parallélisme / Architecture / Système, Toulouse, France, July 2018.

This paper aims in presenting the management of hybrid system using MILP approach.

[NC2] Haddad Marwa, Nicod Jean-Marc, Pera Marie-Cecile Optimisation énergétique des Data-Centers et usages des énergies renouvelables Congrès National de la Recherche des IUT 2017 (CNRIUT 2017), Auxerre, France, May 2017.

This paper aims at presenting an overview about management of renewable energy system.

Research report

[TR1] Haddad Marwa, Nicod Jean-Marc, Christophe Varnier, Pera Marie-Cecile Stand-Alone Renewable Power System Scheduling for a Green Data-Center using Integer Linear Programming, FEMTO-st intitute Research report N: RR–FEMTO-ST–7927, Besançon, France, Mars, 2019. https://hal.archives-ouvertes.fr/hal-02081951v1

This report aims at presenting different management scenarios of management of a hybrid system using MILP approach. This report correspond to the the journal "Stand-Alone Renewable Power System Scheduling for a Green Data-Center using Integer Linear Programming" submitted in Journal of Scheduling

Bibliography

[1] E. McKeogh A. Gonzalez, B. . Gallachóir and K. Lynch. Study of electricity storage technologies and their potential to address wind energy intermittency in ireland. Technical report, Technical report, UCC Sustainable Energy Research Group (SERG), 2004.

[2] Seyed Ali Arefifar, Yasser Abdel-Rady I Mohamed, and Tarek HM El-Fouly. Supply-adequacy-based optimal construction of microgrids in smart distribution systems. *IEEE transactions on smart grid*, 3(3):1491–1502, 2012.

[3] E. Consulting Artelys and G2ELab. Etude sur le potentiel du stockage d'énergies. Technical report, Technical report, ADEME, ATEE, DGCIS, 2013.

[4] S Ashok. Optimised model for community-based hybrid energy system. *Renewable energy*, 32(7):1155–1164, 2007.

[5] Georgia E Asimakopoulou, Aris L Dimeas, and Nikos D Hatziargyriou. Leader-follower strategies for energy management of multi-microgrids. *IEEE transactions on smart grid*, 4(4):1909–1916, 2013.

[6] I Baniasad Askari and M Ameri. Optimal sizing of photovoltaic-battery power systems in a remote region in Kerman, Iran. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 223(5):563–570, 2009.

[7] Chad Augustine, Richard Bain, Jamie Chapman, Paul Denholm, Easan Drury, Douglas G Hall, Eric Lantz, Robert Margolis, Robert Thresher, Debra Sandor, et al. Renewable electricity futures study. volume 2: Renewable electricity generation and storage technologies. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.

[8] Prabodh Bajpai and Vaishalee Dash. Hybrid renewable energy systems for power generation in stand-alone applications: A review. *Renewable and Sustainable Energy Reviews*, 16(5):2926–2939, 2012.

[9] José L Bernal-Agustín, Rodolfo Dufo-López, and David M Rivas-Ascaso. Design of isolated hybrid systems minimizing costs and pollutant emissions. *Renewable Energy*, 31(14):2227–2244, 2006.

[10] Marco Bortolini, Mauro Gamberi, and Alessandro Graziani. Technical and economic design of photovoltaic and battery energy storage system. *Energy Conversion and Management*, 86:81–92, 2014.

[11] MG Bosilovich, R Lucchesi, and M Suarez. MERRA-2: File specification. https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf, 2015.

[12] M Bouzerdoum, A Mellit, and A Massi Pavan. A hybrid model (sarima–svm) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. *Solar Energy*, 98:226–235, 2013.

[13] Stefano Bracco, Gabriele Dentici, and Silvia Siri. Economic and environmental optimization model for the design and the operation of a combined heat and power distributed generation system in an urban area. *Energy*, 55:1014–1024, 2013.

[14] Paul Breeze. Chapter 10 - power system energy storage technologies. In Paul Breeze, editor, *Power Generation Technologies (Third Edition)*, pages 219 – 249. Newnes, third edition edition, 2019.

[15] Peter J. Brockwell and Richard A. Davis. *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1991.

[16] Douglas C Burger, Eric C Peterson, Sompong Paul Olarig, Mark E Shaw, David Kim, and Christian L Belady. Wind-powered data center, November 17 2011. US Patent App. 12/781,445.

[17] Linda Capuano. International energy outlook 2018 (ieo2018). *US Energy Information Administration (EIA): Washington, DC, USA*, 2018:21, 2018.

[18] Giorgio Cau, Daniele Cocco, Mario Petrollese, Søren Knudsen Kær, and Christian Milan. Energy management strategy based on short-term generation scheduling for a renewable microgrid using a hydrogen storage system. *Energy Conversion and Management*, 87:820–831, 2014.

[19] Haisheng Chen, Thang Ngoc Cong, Wei Yang, Chunqing Tan, Yongliang Li, and Yulong Ding. Progress in electrical energy storage system: A critical review. *Progress in natural science*, 19(3):291–312, 2009.

[20] Shaun Coleman and Darren D Garber. Statefulness among clustered satellite platforms, August 1 2017. US Patent 9,722,692.

[21] Gabriele Comodi, Andrea Giantomassi, Marco Severini, Stefano Squartini, Francesco Ferracuti, Alessandro Fonti, Davide Nardi Cesarini, Matteo Morodo, and Fabio Polonara. Multi-apartment residential microgrid with electrical and thermal storage devices: Experimental analysis and simulation of energy management strategies. *Applied Energy*, 137:854–866, 2015.

[22] John Conti, Paul Holtberg, Jim Diefenderfer, Angelina LaRose, James T Turnure, and Lynn Westfall. International energy outlook 2016 with projections to 2040. Technical report, USDOE Energy Information Administration (EIA), Washington, DC (United States . . . , 2016.

[23] Georges Da Costa. Heterogeneity: The key to achieve power-proportional computing. In *2013 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing*, pages 656–662. IEEE, 2013.

[24] Hanane Dagdougui, Ahmed Ouammi, and Roberto Sacile. Modelling and control of hydrogen and energy flows in a network of green hydrogen refuelling stations powered by mixed renewable energy systems. *International Journal of Hydrogen Energy*, 37(6):5360–5371, 2012.

[25] Mathieu David, Faly H Ramahatana Andriamasomanana, and Olivier Liandrat. Spatial and temporal variability of pv output in an insular grid: Case of reunion island. *Energy Procedia*, 57:1275–1282, 2014.

[26] Ali Demirbas. Present and future transportation fuels. *Energy Sources, Part A*, 30(16):1473–1483, 2008.

[27] Ayhan Demirbas. Potential applications of renewable energy sources, biomass combustion problems in boiler power systems and combustion related environmental issues. *Progress in energy and combustion science*, 31(2):171–192, 2005.

[28] Wei Deng, Fangming Liu, Hai Jin, and Xiaofei Liao. Online control of datacenter power supply under uncertain demand and renewable energy. In *Communications (ICC), 2013 IEEE International Conference on*, pages 4228–4232. IEEE, 2013.

[29] Hadja Maïmouna Diagne. *Gestion intelligente du réseau électrique réunionnais. Prévision de la ressource solaire en milieu insulaire*. PhD thesis, Université de la Réunion, 2015.

[30] Andrew Dicks and David Anthony James Rand. *Fuel cell systems explained*. Wiley Online Library, 2018.

[31] Weiqiang Dong, Yanjun Li, and Ji Xiang. Optimal sizing of a stand-alone hybrid power system based on battery/hydrogen with an improved ant colony optimization. *Energies*, 9(10):785, 2016.

[32] Samah Elsayed Sonia Rueda Iana Arkhipova drian Whiteman, Javier Esparrago and Gerardo Escamilla. Renewable energy statistics 2019. Technical report, The International Renewable Energy Agency (IRENA), 2019.

[33] Bob Dudley et al. Bp statistical review of world energy. *London, UK*, 2012.

[34] Rodolfo Dufo-Lopez and José L Bernal-Agustín. Design and control strategies of PV-diesel systems using genetic algorithms. *Solar energy*, 79(1):33–46, 2005.

[35] Economic Development Blog. MidAmerican Energy to Invest $1.9B for Iowa Wind Generation, 2013. Available: https://goo.gl/dcxBQJ (Accessed Nov. 2018).

[36] Ottmar Edenhofer, Ramón Pichs-Madruga, Youba Sokona, Kristin Seyboth, Susanne Kadner, Timm Zwickel, Patrick Eickemeier, Gerrit Hansen, Steffen Schlömer, Christoph von Stechow, et al. *Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change*. Cambridge University Press, 2011.

[37] Michael W Ellis, Michael R Von Spakovsky, and Douglas J Nelson. Fuel cell systems: efficient, flexible energy conversion for the 21st century. *Proceedings of the IEEE*, 89(12):1808–1818, 2001.

[38] Aerowatt and edf-sei. enerstock: cahier des charges du stockage, 2011.

[39] Mustafa Engin and Dilşad Engin. Sizing PV-wind hybrid energy system for lighting. *International Journal of Development and Sustainability*, 1(2):85–98, 2012.

[40] Tomoya Enokido and Makoto Takizawa. An integrated power consumption model for distributed systems. *IEEE Transactions on Industrial Electronics*, 60(2):824–836, 2012.

[41] Aitor Etxeberria, Ionel Vechiu, Haritza Camblong, and Jean-Michel Vinassa. Hybrid energy storage systems for renewable energy sources integration in microgrids: A review. In *2010 Conference Proceedings IPEC*, pages 532–537. IEEE, 2010.

[42] Cédric FERLAT. Le stockage hydraulique gravitaire et la production hydroélectrique sur les pentes réunionnaises–etude du stockage de l'énergie photovoltaïque, 2009.

[43] Lee Fingersh, Dave Simms, Maureen Hand, Dave Jager, Jason Cotrell, Mike Robinson, Scott Schreck, and Scott M Larwood. Wind tunnel testing of NREL's unsteady aerodynamics experiment. In *20th ASME Wind Energy Symposium*, 2001.

[44] Christian Francq and Antony Gautier. Large sample properties of parameter least squares estimates for time-varying ARMA models. *J. Time Ser. Anal.*, 25(5):765–783, 2004.

[45] Raquel S Garcia and Daniel Weisser. A wind–diesel system with hydrogen storage: Joint optimisation of design and dispatch. *Renewable energy*, 31(14):2296–2320, 2006.

[46] Íñigo Goiri, Ryan Beauchea, Kien Le, Thu D Nguyen, Md E Haque, Jordi Guitart, Jordi Torres, and Ricardo Bianchini. Greenslot: scheduling energy consumption in green datacenters. In *High Performance Computing, Networking, Storage and Analysis (SC), 2011 International Conference for*, pages 1–11. IEEE, 2011.

[47] Inigo Goiri, Md E. Haque, Kien Le, Ryan Beauchea, Thu D. Nguyen, Jordi Guitart, Jordi Torres, and Ricardo Bianchini. Matching renewable energy supply and demand in green datacenters. *Ad Hoc Networks*, 25:520 – 534, 2015.

[48] Green data net project, 2016. http://www.greendatanet-project.eu/home.html.

[49] Marwa Haddad, Jean-Marc Nicod, and Marie-Cecile Marion-Pera. Hydrogen infrastructure: Data-center supply-refueling station synergy. In *IEEE Vehicle Power and Propulsion Conference (VPPC'2017)*, 12 2017.

[50] Jamal Hassan. Arima and regression models for prediction of daily and monthly clearness index. *Renewable Energy*, 68:421–427, 2014.

[51] Keith Hipel and A. Ian McLeod. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam, 1994.

[52] U Hoelzle and L Barroso. The datacenter as a computer. *Morgan and Claypool*, 2009.

[53] John P Holdren, Kirk R Smith, Tord Kjellstrom, David Streets, and Xiaodong Wang. World energy assessment: energy and the challenge of sustainability. *New York: United Nations Development Programme*, pages 61–110, 2000.

[54] L Hontoria, J Aguilera, and P Zufiria. Generation of hourly irradiation synthetic series using the neural network multilayer perceptron. *Solar Energy*, 72(5):441–446, 2002.

[55] L Hontoria, J Aguilera, and P Zufiria. A new approach for sizing stand alone photovoltaic systems based in neural networks. *Solar Energy*, 78(2):313–319, 2005.

[56] Ramin Hosseinalizadeh, Hamed Shakouri, Mohsen Sadegh Amalnick, and Peyman Taghipour. Economic sizing of a hybrid (pv–wt–fc) renewable energy system (hres) for stand-alone usages by an optimization-simulation model: case study of iran. *Renewable and Sustainable Energy Reviews*, 54:139–150, 2016.

[57] Peng Hou, Weihao Hu, Baohua Zhang, Mohsen Soltani, Cong Chen, and Zhe Chen. Optimised power dispatch strategy for offshore wind farms. *IET Renewable Power Generation*, 10(3):399–409, 2016.

[58] Greenpeace Inc. Clicking clean: how companies are creating the green internet, 2014. Available: https://goo.gl/7XJsh5 (Accessed Nov. 2018).

[59] Quanyuan Jiang, Meidong Xue, and Guangchao Geng. Energy management of microgrid in grid-connected and stand-alone modes. *IEEE transactions on power systems*, 28(3):3380–3389, 2013.

[60] JK Kaldellis, D Zafirakis, and E Kondili. Optimum autonomous stand-alone photovoltaic system design on the basis of energy pay-back analysis. *Energy*, 34(9):1187–1198, 2009.

[61] A. Kassab, J. M. Nicod, L. Philippe, and V. Rehn-Sonigo. Scheduling independent tasks in parallel under power constraints. In *46th International Conference on Parallel Processing (ICPP)*, pages 543–552, Aug 2017.

[62] Ayham Kassab, Jean-Marc Nicod, Laurent Philippe, and Veronika Rehn-Sonigo. Assessing the use of genetic algorithms to schedule independent tasks under power constraints. In *2018 International Conference on High Performance Computing & Simulation (HPCS)*, pages 252–259. IEEE, 2018.

[63] YA Katsigiannis, PS Georgilakis, and ES Karapidakis. Multiobjective genetic algorithm solution to the optimum economic and environmental performance problem of small autonomous hybrid power systems with renewables. *IET Renewable Power Generation*, 4(5):404–419, 2010.

[64] Rajesh G Kavasseri and Krithika Seetharaman. Day-ahead wind speed forecasting using f-arima models. *Renewable Energy*, 34(5):1388–1393, 2009.

[65] Amin Khodaei, Shay Bahramirad, and Mohammad Shahidehpour. Microgrid planning under uncertainty. *IEEE Transactions on Power Systems*, 30(5):2417–2425, 2014.

[66] Jonathan Koomey. Growth in data center electricity use 2005 to 2010. *A report by Analytical Press, completed at the request of The New York Times*, 9, 2011.

[67] V Kostylev, A Pavlovski, et al. Solar power forecasting performance–towards industry standards. In *1st international workshop on the integration of solar power into power systems, Aarhus, Denmark*, 2011.

[68] Phillip Oliver Kriett and Matteo Salani. Optimal control of a residential microgrid. *Energy*, 42(1):321–330, 2012.

[69] James Larminie, Andrew Dicks, and Maurice S McDonald. *Fuel cell systems explained*, volume 2. J. Wiley Chichester, UK, 2003.

[70] Alois Lechner. Pumped storage for the future. *hydronews*, 6(3), 2018.

[71] D. Lefebvre. Enerstock - cahier des charges du stockage. technical report. Technical report, Aerowatt - EDF-SEI, 2011.

[72] Bei Li, Robin Roche, and Abdellatif Miraoui. Microgrid sizing with combined evolutionary algorithm and MILP unit commitment. *Applied energy*, 188:547–562, 2017.

[73] Bei Li, Robin Roche, Damien Paire, and Abdellatif Miraoui. Sizing of a stand-alone microgrid considering electric power, cooling/heating, hydrogen loads and hydrogen storage degradation. *Applied Energy*, 205:1244–1259, 2017.

[74] Chun-Hua Li, Xin-Jian Zhu, Guang-Yi Cao, Sheng Sui, and Ming-Ruo Hu. Dynamic modeling and sizing optimization of stand-alone photovoltaic power systems using hybrid energy storage technology. *Renewable energy*, 34(3):815–826, 2009.

[75] Y. Li, A. C. Orgerie, and J. M. Menaud. Balancing the use of batteries and opportunistic scheduling policies for maximizing renewable energy consumption in a cloud data center. In *2017 25th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, pages 408–415, March 2017.

[76] Yongxin Li, Bing Xue, and Xueyi He. Catalytic synthesis of ethylbenzene by alkylation of benzene with diethyl carbonate over hzsm-5. *Catalysis Communications*, 10(5):702–707, 2009.

[77] Zhenhua Liu, Yuan Chen, Cullen Bash, Adam Wierman, Daniel Gmach, Zhikui Wang, Manish Marwah, and Chris Hyser. Renewable and cooling aware workload management for sustainable data centers. *ACM SIGMETRICS Performance Evaluation Review*, 40(1):175–186, 2012.

[78] Elke Lorenz, Annette Hammer, Detlev Heinemann, et al. Short term forecasting of solar radiation based on satellite data. In *EUROSUN2004 (ISES Europe Solar Congress)*, pages 841–848, 2004.

[79] Bo Lu and Mohammad Shahidehpour. Short-term scheduling of battery in a grid-connected pv/battery system. *IEEE Transactions on Power Systems*, 20(2):1053–1061, 2005.

[80] Ricardo Luna-Rubio, Mario Trejo-Perea, D Vargas-Vázquez, and GJ Ríos-Moreno. Optimal sizing of renewable hybrids energy systems: A review of methodologies. *Solar Energy*, 86(4):1077–1088, 2012.

[81] Akbar Maleki and Alireza Askarzadeh. Comparative study of artificial intelligence techniques for sizing of a hydrogen-based stand-alone photovoltaic/wind hybrid system. *international journal of hydrogen energy*, 39(19):9973–9984, 2014.

[82] Jennifer Mankoff, Robin Kravets, and Eli Blevis. Some computer science issues in creating a sustainable world. *Computer*, 41(8):102–105, August 2008.

[83] John Markoff. Microsoft plumbs oceans depths to test underwater data center. *The New York Times*, 2016.

[84] Gilbert M Masters. *Renewable and efficient electric power systems*. John Wiley & Sons, 2013.

[85] Eugenia D Mehleri, Haralambos Sarimveis, Nikolaos C Markatos, and Lazaros G Papageorgiou. A mathematical programming approach for optimal design of distributed energy systems at the neighbourhood level. *Energy*, 44(1):96–104, 2012.

[86] A Mellit, M Benghanem, A Hadj Arab, and A Guessoum. An adaptive artificial neural network model for sizing stand-alone photovoltaic systems: application for isolated sites in algeria. *Renewable Energy*, 30(10):1501–1524, 2005.

[87] A Mellit, M Benghanem, A Hadj Arab, and A Guessoum. A simplified model for generating sequences of global solar radiation data for isolated sites: Using artificial neural network and a library of markov transition matrices approach. *Solar Energy*, 79(5):469–482, 2005.

[88] Adel Mellit, M Benghanem, and Soteris A Kalogirou. Modeling and simulation of a stand-alone photovoltaic system using an adaptive artificial neural network: Proposition for a new sizing procedure. *Renewable energy*, 32(2):285–313, 2007.

[89] Rich Miller. Huge solar array will support nj data center. *datacenter-knowledge*, 2011. http://www.datacenterknowledge.com/archives/2011/06/14/huge-solar-array-will-support-nj-data-center/.

[90] M Mohammadi, SH Hosseinian, and GB Gharehpetian. Optimization of hybrid solar energy sources/wind turbine systems integrated to utility grids as microgrid (mg) under pool/bilateral/hybrid electricity market using pso. *Solar energy*, 86(1):112–125, 2012.

[91] Sirus Mohammadi, Babak Mozafari, Soodabeh Solimani, and Taher Niknam. An adaptive modified firefly optimisation algorithm based on hong's point estimate method to optimal operation management in a microgrid with consideration of uncertainties. *Energy*, 51:339–348, 2013.

[92] C Monteiro, H Keko, R Bessa, V Miranda, A Botterud, J Wang, G Conzelmann, et al. A quick guide to wind power forecating: state-of-the-art 2009. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States), 2009.

[93] A Moschetto, G Giaquinta, and G Tina. Modelling of integrated renewable energy systems supported by hydrogen storage. In *Power Tech, 2007 IEEE Lausanne*, pages 2088–2092. IEEE, 2007.

[94] Bernard MULTON and Jacques Ruer. Stocker l'électricité: oui, c'est indispensable et c'est possible. pourquoi, où, comment? *Publication ECRIN en contribution au débat national sur l'énergie*, 2003.

[95] Nirmal-Kumar C Nair and Niraj Garimella. Battery energy storage systems: Assessment for small-scale renewable energy integration. *Energy and Buildings*, 42(11):2124–2130, 2010.

[96] NREL. Hybrid optimization model for electric renewable energy, 2009.

[97] D. O'Sullivan. Data centre redundancy levels. Technical report, EATON, March 2016.

[98] G. Parise and L. Parise. Electrical distribution for a reliable data center. In *2012 IEEE Industry Applications Society Annual Meeting*, pages 1–5, Oct 2012.

[99] Michael K Patterson. The effect of data center temperature on energy efficiency. In *2008 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pages 1167–1174. IEEE, 2008.

[100] Birgit Penzenstadler, Bill Tomlinson, Eric Baumer, Marcel Pufal, Ankita Raturi, Debra Richardson, Baki Cakici, Ruzanna Chitchyan, Georges Da Costa, Lynn Dombrowski, Malin Picha Edwardsson, Elina Eriksson, Xavier Franch, Gillian R. Hayes, Christina Herzog, Wolfgang Lohmann, Martin Mahaux, Alistair Mavin, Melissa Mazmanian, Sahand Nayebaziz, Juliet Norton, Daniel Pargman, Donald J. Patterson, Jean-Marc Pierson, Kristin Roher, M. Six Silberman, Kevin Simonson, Andrew W. Torrance, and André Van Der Hoek. ICT4S 2029: What Will Be The Systems Supporting Sustainability in 15 Years (regular paper). In *ICT for Sustainability (ICT4S), Stockholm, 24/08/2014-27/08/2014*, pages 30–39, http://www.atlantis-press.com, 2014. Atlantis Press.

[101] Jacques Percebois and Claude Mandil. Rapport énergies 2050. *Le nucléaire un an après Fukushima*, (1):13, 2012.

[102] J Pestourie. Les enr intermittentes. cas des réseaux insulaires. Technical report, Technical report, EDF R&D, 2012.

[103] J. Pierson, G. Baudic, S. Caux, B. Celik, G. Da Costa, L. Grange, M. Haddad, J. Lecuivre, J. Nicod, L. Philippe, V. Rehn-Sonigo, R. Roche, G. Rostirolla, A. Sayah, P. Stolf, M. Thi, and C. Varnier. Datazero: Datacenter with zero emission and robust management using renewable energy. *IEEE Access*, 7:103209–103230, 2019.

[104] Jack Pouchet. Biomass-powered data centers: Next step for green it?, 2012.

[105] Thomas Pregger, Daniela Graf, Wolfram Krewitt, Christian Sattler, Martin Roeb, and Stephan Möller. Prospects of solar thermal hydrogen production processes. *International journal of hydrogen energy*, 34(10):4256–4267, 2009.

[106] C Protogeropoulos, BJ Brinkworth, and RH Marshall. Sizing and techno-economical optimization for hybrid solar photovoltaic/wind power systems with battery storage. *International Journal of Energy Research*, 21(6):465–479, 1997.

[107] S. Qi, W. Sun, and Y. Wu. Comparative analysis on different architectures of power supply system for data center and telecom center. In *2017 IEEE International Telecommunications Energy Conference (INTELEC)*, pages 26–29, Oct 2017.

[108] Daniel Quiggin, Sarah Cornell, Michael Tierney, and Richard Buswell. A simulation and optimisation study: Towards a decentralised microgrid, using real world fluctuation data. *Energy*, 41(1):549–559, 2012.

[109] Singiresu S Rao. *Engineering optimization: theory and practice*. John Wiley & Sons, 2009.

[110] Neil Rasmussen. Determining total cost of ownership for data center and network room infrastructure. *Relatório técnico, Schneider Electric, Paris*, 8, 2011.

[111] Neil Rasmussen and Wendy Torell. Comparing data center power distribution architectures. *American Power Conversion, Tech. White Paper*, 129, 2011.

[112] Renewit project, 2016. http://www.renewit-project.eu/.

[113] Gaël Robin, Marie Ruellan, Bernard Multon, Hamid Ben Ahmed, and Pierre-Yves Glorennec. Solutions de stockage de l'énergie pour les systèmes de production intermittente d'électricité renouvelable. In *Colloque SeaTechWeek 2004 (Semaine Internat. des Technologies de la Mer)*, page 9p, 2004.

[114] Alan Ruddell. Investigation on storage technologies for intermittent renewable energies: Evaluation and recommended r&d strategy. *Storage Technology report ST6: Flywheel, Chilton Didcot www. itpower. co. uk/investire/pdfs/flywheelrep. pdf*, 2003.

[115] Mohammad Masih Sediqi, Masahiro Furukakoi, Mohammed E Lotfy, Atsushi Yona, and Tomonobu Senjyu. Optimal economical sizing of grid-connected hybrid renewable energy system. *Journal of Energy and Power Engineering*, 11(4):244–53, 2017.

[116] Manajit Sengupta, Yu Xie, Anthony Lopez, Aron Habte, Galen Maclaurin, and James Shelby. The national solar radiation data base (NSRDB). *Renewable and Sustainable Energy Reviews*, 89:51–60, 2018.

[117] Jun-Hai Shi, Xin-Jian Zhu, and Guang-Yi Cao. Design and techno-economical optimization for stand-alone hybrid power systems with multi-objective evolutionary algorithms. *International Journal of Energy Research*, 31(3):315–328, 2007.

[118] Nirendra Lal Shrestha, Noah Pflugradt, Thorsten Urbaneck, Angel Carrera, Eduard Oró, Albert Garcia, Hans Trapman, Gilbert de Nijis, Joris van Dorp (DEERNS), and Mario Macías (BSC). Deliverable 4.5 catalogue of advanced technical concepts for net zero energy data centres. Technical report, RenewIT, 30 Sept. 2015.

[119] Javier Silvente, Georgios M Kopanos, Efstratios N Pistikopoulos, and Antonio Espuña. A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids. *Applied Energy*, 155:485–501, 2015.

[120] Sunanda Sinha and SS Chandel. Review of software tools for hybrid renewable energy systems. *Renewable and Sustainable Energy Reviews*, 32:192–205, 2014.

[121] Sunanda Sinha and SS Chandel. Review of recent trends in optimization techniques for solar photovoltaic–wind based hybrid energy systems. *Renewable and Sustainable Energy Reviews*, 50:755–769, 2015.

[122] Jayantha Siriwardana, Saliya Jayasekara, and Saman K Halgamuge. Potential of air-side economizers for data center cooling: A case study for key australian cities. *Applied Energy*, 104:207–219, 2013.

[123] Petr Stluka, Datta Godbole, and Tariq Samad. Energy management for buildings and microgrids. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 5150–5157. IEEE, 2011.

[124] SR Tewfik. Biomass utilization facilities and biomass processing technologies. *Energy Education Science and Technology*, 14:1–19, 2004.

[125] G Tina and S Gagliano. Probabilistic analysis of weather data for a hybrid solar/wind energy system. *International Journal of Energy Research*, 35(3):221–232, 2011.

[126] My Ton, Brian Fortenbery, and William Tschudi. Dc power for improved data center efficiency. *Lawrence Berkeley National Laboratory*, 2008.

[127] Huan-Liang Tsai, Ci-Siang Tu, Yi-Jie Su, et al. Development of generalized photovoltaic model using matlab/simulink. In *Proceedings of the world congress on Engineering and computer science*, volume 2008, pages 1–6. San Francisco, USA, 2008.

[128] Rahul Urgaonkar, Bhuvan Urgaonkar, Michael J Neely, and Anand Sivasubramaniam. Optimal power cost management using stored energy in data centers. In *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, pages 221–232. ACM, 2011.

[129] M Uzunoglu, OC Onar, and MS Alam. Modeling, control and simulation of a pv/fc/uc based hybrid power generation system for stand-alone applications. *Renewable energy*, 34(3):509–520, 2009.

[130] Thong Vu Van, Mark Norton, Chavdar Ivanov, Marko Delimar, Nikos Hatziargyriou, Jon Stromsather, Antonio Iliceto, Carlos Llanos, and Patrick Panciatici. Organic growth: toward a holistic approach to european research and innovation. *IEEE Power and Energy Magazine*, 13(1):30–37, 2015.

[131] Micha vor dem Berge, Jochen Buchholz, Leandro Cupertino, Georges Da Costa, Andrew Donoghue, Georgina Gallizo, Mateusz Jarus, Lara Lopez, Ariel Oleksiak, Enric Pages, et al. Coolemall: Models and tools for planning and operating energy efficient data centres. In *Handbook on Data Centers*, pages 191–245. Springer, 2015.

[132] Cyril Voyant, Marc Muselli, Christophe Paoli, and Marie-Laure Nivet. Numerical weather prediction (nwp) and hybrid arma/ann model to predict global radiation. *Energy*, 39(1):341–355, 2012.

[133] M Mitchell Waldrop. Data center in a box. *Scientific American*, 297(2):90–93, 2007.

[134] Lingfeng Wang and Chanan Singh. Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm. *IEEE Transactions on Energy Conversion*, 24(1):163–172, 2009.

[135] Joakim Widén, Nicole Carpman, Valeria Castellucci, David Lingfors, Jon Olauson, Flore Remouit, Mikael Bergkvist, Mårten Grabbe, and Rafael Waters. Variability assessment and forecasting of renewables: A review for solar, wind, wave and tidal resources. *Renewable and Sustainable Energy Reviews*, 44:356–375, 2015.

[136] F. Yang and A. A. Chien. Large-scale and extreme-scale computing with stranded green power: Opportunities and costs. *IEEE Transactions on Parallel and Distributed Systems*, 29(5):1103–1116, May 2018.

[137] Hongxing Yang, Zhou Wei, and Lou Chengzhi. Optimal design and techno-economic analysis of a hybrid solar–wind power generation system. *Applied Energy*, 86(2):163–169, 2009.

[138] Hongxing Yang, Wei Zhou, Lin Lu, and Zhaohong Fang. Optimal sizing method for stand-alone hybrid solar–wind system with LPSP technology by using genetic algorithm. *Solar energy*, 82(4):354–367, 2008.

[139] HX Yang, L Lu, and J Burnett. Weather data and probability analysis of hybrid photovoltaic–wind power generation systems in hong kong. *Renewable Energy*, 28(11):1813–1824, 2003.

[140] Global Energy Statistical Yearbook. Electricity consumption. *Enerdata.[access: yearbook. enerdata. net]*, 2019.

[141] Yorksolar. off grid solar installations, 2011. http://yorksolar.ca/off-the-grid.html.

[142] Alireza Zakariazadeh, Shahram Jadid, and Pierluigi Siano. Smart microgrid energy and reserve scheduling with demand response using stochastic optimization. *International Journal of Electrical Power & Energy Systems*, 63:523–533, 2014.

[143] J Zeng, M Li, JF Liu, J Wu, and HW Ngan. Operational optimization of a stand-alone hybrid renewable energy generation system based on an improved genetic algorithm. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–6. IEEE, 2010.

[144] Di Zhang, Nilay Shah, and Lazaros G Papageorgiou. Efficient energy consumption and operation management in a smart building with microgrid. *Energy Conversion and Management*, 74:209–222, 2013.

[145] Hainan Zhang, Shuangquan Shao, Hongbo Xu, Huiming Zou, and Changqing Tian. Free cooling of data centers: A review. *Renewable and Sustainable Energy Reviews*, 35:171–182, 2014.

[146] Binyan Zhao, Yi Shi, Xiaodai Dong, Wenpeng Luan, and Jens Bornemann. Short-term operation scheduling in renewable-powered microgrids: A duality-based approach. *IEEE Transactions on Sustainable Energy*, 5(1):209–217, 2013.

[147] Bahman Zohuri. *Hybrid Energy Systems: Driving Reliable Renewable Sources of Energy Storage*. Springer, 2017.

Appendix A

Power load determination

The obtained of the power load determination tool is summarized in the following table:

Table A.1: Results of the power load determination tool

Year	*Nwt*	*Npv*	*Battery capacity*	*Hydrogene size*	*Demand*
2004	0	17487	7840.38	15118.42	499.98
	1	11505	6666.20	9656.63	500.00
	2	7075	4844.50	6748.92	500.00
	3	3382	5977.39	4378.75	500.01
	4	148	8063.26	6436.34	500.02
	5	0	10124.00	8168.31	617.86
2005	0	18269	8027.56	15511.10	500.02
	1	11720	6927.48	10301.83	499.98
	2	7011	5546.82	8026.92	499.99
	3	3050	5649.51	9581.05	499.98
	4	0	8603.93	12241.93	517.32
2006	0	17865	8199.86	12991.43	499.99
	1	11840	6963.80	7959.60	500.01
	2	7562	5112.36	5497.27	499.99
	3	3983	5526.27	6745.69	500.03
	4	828	7507.09	10702.41	500.03
	5	0	9165.26	13994.54	585.83
2007	0	17249	8572.21	13538.50	500.01
	1	12179	7310.12	9720.98	500.02
	2	8456	6786.12	8672.29	500.02
	3	5329	6547.33	8689.68	499.99
	4	2475	6346.00	10991.14	500.01
	5	0	7259.95	13672.79	508.94
2008	0	17585	7911.66	14540.78	500.02
	1	12559	7204.35	10929.36	500.01
	2	8712	6775.97	9701.63	500.00
	3	5560	6502.30	8884.50	499.97
	4	2690	6844.85	9079.32	500.02
	5	0	8749.28	10115.18	500.42

Table A.2: Results of the power load determination tool

Year	Nwt	Npv	Battery capacity	Hydrogene size	Demand
2009	0	17675	8055.71	14107.47	500.03
	1	12494	7620.75	9587.30	500.02
	2	8744	6928.33	8426.53	500.03
	3	5631	7036.50	10345.07	500.00
	4	2772	6442.56	12765.82	500.03
	5	80	7622.88	15887.82	499.99
	6	0	9174.98	19063.57	596.45
2010	0	18084	8065.37	15892.14	500.02
	1	13663	7533.78	12111.10	499.98
	2	10330	7141.64	12570.24	500.00
	3	7669	7117.26	13618.17	500.00
	4	5185	7370.03	15550.29	500.00
	5	2866	8054.65	17645.36	499.99
	6	659	8796.60	20624.71	500.03
	7	0	10019.60	23974.16	555.60
2011	0	17313	8056.63	13251.32	500.02
	1	12440	7473.14	10852.20	500.01
	2	8741	6790.08	10025.20	500.00
	3	5642	5568.30	10998.12	500.02
	4	2796	5764.14	13654.40	500.02
	5	117	7952.21	16487.90	499.98
	6	0	9610.65	19760.49	594.78
2012	0	17777	7930.42	13964.71	500.00
	1	12916	7507.83	10559.48	499.98
	2	9196	7466.17	8855.29	500.00
	3	6154	7587.62	8735.73	499.99
	4	3375	7690.78	9706.44	500.03
	5	780	7779.74	10730.94	500.02
	6	0	9285.35	13003.42	565.71

Appendix B

Forecasting Methodology

Before presenting the whole methodology, the type of data as well as their locations used are those of the chapters of this manuscript. So, here we dispose of two types of data: solar radiation and wind speed. The latter could be obtained from various databases online such as the national solar data base (NSRDB) [116], the Modern-Era Retrospective analysis for Research and Applications (MERRA2) [11], the wind prospector from the National Renewable Energy Laboratory (NREL) [43]. In our case, the data are obtained from NSRDB AND NREL. Recall that the aim of this chapter is to propose a statistical approach for wind and solar forecasting. For that purpose, based on a review and results obtained by different researchers mentioning the accuracy of the ARIMA model [12, 50, 64, 132], we have selected the SARIMA model [15, 51]. In order to verify the robustness of the SARIMA approach on our application, we will apply the methodology on two distinct locations having different characteristics.

B.0.1/ SARIMA model

ARIMA is a statisctical approach widely used in today's world since the evolution of sophisticated statistical software package. ARIMA has four major steps in model building-Identification, Estimation, Diagnostics & Forecast. Then, the general scheme for ARIMA model is translated by:

1. Identification of the model structure;
2. Application of autocorrelation function (ACF) and partial autocorrelation function (PACF) in order to identify the orders of the ARMA model. The parameters of the model are estimated by a maximum likelihood (ML) function;
3. Testing the goodness of fit on the estimated model residuals;
4. Using the estimated model for forecasting.

ARIMA model uses the historic data and decomposes it into autoregressive (AR), Integrated (I) indicates linear trends or polynomial trend and Moving Average (MA) indicates weighted moving average over past errors. Therefore, it has three model parameters AR(p), I(d) and MA(q) all combined to form ARIMA(p, d, q) model where:

- p = order of AR;

- q = order of MA;
- d = order of I (differencing).

The multiplicatif Seasonal ARIMA model namely SARIMA is actually a variation of the classical ARIMA model. In order to take into account the seasonal effect of the irradiation and the wind speed, this model is generally written as SARIMA(p,d,q)(P,D,Q) where, as in the ARIMA model, p, d, q and P, D, Q are non-negative integers that refer to the polynomial order of the AR, I, MA parts of the non-seasonal and seasonal components of the model, respectively. Mathematically, the SARIMA model is defined as in (B.1)

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D x_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{B.1}$$

Where: x_t is the forecast variable (i.e., solar radiation), $\phi_p(B)$ is the regular AR polynomial of order $p()$, $\theta_q(B)$ is the regular MA polynomial of order $q()$, $\Phi_P(B^s)$ is the seasonal AR polynomial of order $P()$, $\Theta_Q(B^s)$ is the seasonal MA polynomial of order Q, ∇^d is the differentiating operator that eliminate the non-seasonal non-stationarity, ∇_s^D is the seasonal differentiating operator that eliminate the seasonal non-stationarity, B is the backshift operator, which shift one point in time the observation x_t (i.e., $B^k(x_t) = x_{t-k}$) and finally ε_t follows a white noise process and s defines the seasonal period. These polynomials are described mathematically in Equations (B.2):

$$\begin{aligned} \theta_q(B) &= 1 - \sum_{i=1}^{q}\theta_i B^i & \Theta_Q(B^s) &= 1 - \sum_{i=1}^{Q}\Theta_i B^{s,i} \\ \phi_p(B) &= 1 - \sum_{i=1}^{p}\phi_i B^i & \Phi_P(B^s) &= 1 - \sum_{i=1}^{P}\Phi_i B^{s,i} \\ \nabla^d &= (1-B)^d & \nabla_s^d &= (1-B^s)^D \end{aligned} \tag{B.2}$$

In order to get the model that fits the best the data, the Akaike Information Criterion (AIC) is a statistic measure to compare them. In fact, the AIC rewards models for a good fit and penalize others for complexity. It could be written as:

$$AIC = 2k + \ln\left(\frac{RSS}{n}\right) \tag{B.3}$$

with k the number of free parameters, n the total number of observations equal to 468 and RSS is the residual sum of squares.

Finally, using the obtained valid model, one can proceed to the forecasting of the wished period.

B.0.2/ Evaluation of the forecasting performance

The forecasting model is constructed on time series of solar radiation and wind speed for a duration of 9 years weekly (which means 468 values: average weekly values) Once the models are formulated, they are used to forecast wind speed and solar radiation for the last two years. Afterwards, the averages of the statistics for the 2 years forecasting results are computed to analyze the models' accuracy. Several measurement statistics can be used to examine the forecast accuracy of different models. Mean absolute percentage error (MAPE) is used very often to evaluate the performance of the forecasting model. The above-mentioned statistical quantities are computed as in (B.4):

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{x_t - \hat{x}_t}{x_t} \right| \tag{B.4}$$

where $\hat{x}_t$ is the forecast value.

B.1/ Results and discussions

The data representing the solar radiation and wind speed were measured on an hourly scale from January 2004 till December 2012 (more than 6 years) in two different location. To be more precise, the endogenous data of the solar radiation and wind speed time series were measured at Chicago (Latitude: 41.810539, Longitude: -87.643127, Time Zone: -6) and at Los Angeles (Latitude: 34.57, Longitude: -118.02, Time Zone: -8). Then, in order to obtain weekly values, we calculated averages per groups of 168 values (168 hours per week). Finally, we have obtained time series of 52 value per year, i.e., 468 values during the nine years. Figures B.1 and B.2 showed this distribution respectively for solar radiation and wind speed in Los Angeles. The first nine years have been used to setup our models and the last two years to test them. The model has been implemented using R programming language.

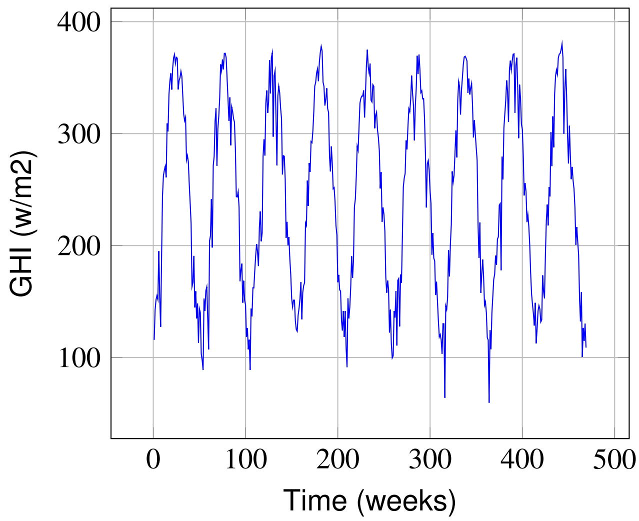


Figure B.1: Weekly solar radiation distribution in Los Angeles

B.1.1/ Models validation

Based on Figure B.1, the measured solar radiation from 2004 till 2012 is quite seasonal. In fact, the data starts from the first week of January till the last week of December. Each year, the pic of solar radiation is in July that corresponds to the summer season where days are quite long. Contrariwise, the lowest values are obtained in December or in January.

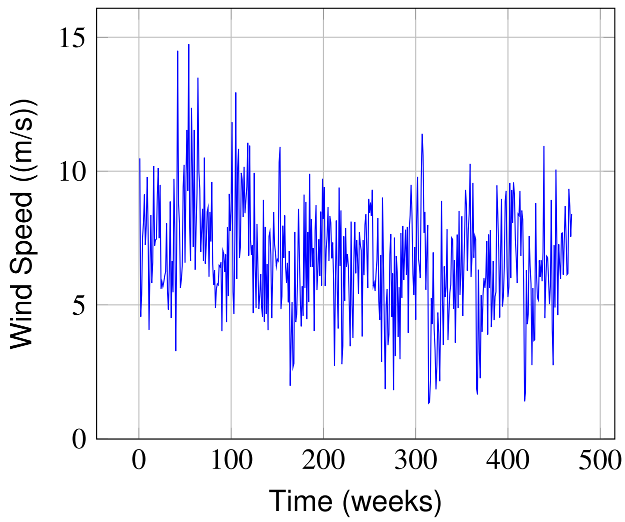


Figure B.2: Weekly wind speed distribution in Los Angeles

This period matches with the winter where days are short. Thus, the solar distribution is intrinsically seasonal and periodic which validates the choice of the SARIMA model.

In Figure B.2, the data also starts from the first week of January till the last week of December. Moreover, it shows a random distribution where data varies from 3 m/s till 14.8 m/s. This series presents a seasonality that could be well seen especially starting from the week 150. The wind speed is quite low in the winter and increases with the oncoming of summer corresponding to the thermal hot wind of Los Angeles. Nevertheless, the wind can vary from one year to another so we cannot confirm the periodicity.

Table B.1: Comparison of the statistic criterion AIC for wind speed in both Chicago and Los Angeles

DATA	***SARIMA Configurations***	***AIC***
Chicago	SARIMA(21,0,21)(1,1,0)	1441,393
	SARIMA(11,0,14)(1,1,1)	1359,254
	SARIMA(11,0,14)(0,1,1)	1358,813
	SARIMA(18,0,18)(0,1,0)	1498,97
Los Angeles	SARIMA(9,0,19)(1,1,0)	1886,61
	SARIMA(6,0,6)(1,1,1)	1844,582
	SARIMA(6,0,6)(0,1,1)	1839,007
	SARIMA(9,0,0)(0,1,0)	2006,21

In order to obtain the model that fits the best the data, different configurations of SARIMA have been applied on the distributions for the two locations. In each set, 4 seasonal configurations have been applied such as the seasonal polynomial AR and MA respectively $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are set as explained in Table B.1

Based on results given in Tables B.1 and B.2 in the two different cities, with completely different characteristics, one can see that the best AIC obtained is the one of the model configuration SARIMA(p,d,q)$(0,1,1)$ for both solar radiation and wind speed data. For instance, The SARIMA model$(6,0,6)(0,1,1)$ is written during a period of $s = 52$ as in Equation (B.5)

Table B.2: Comparison of the statistic criterion AIC for solar radiation in both Chicago and Los Angeles

DATA	***SARIMA Configuration***	***AIC***
Chicago	SARIMA(10,0,9)(1,1,0)	4162,49
	SARIMA(4,0,18)(1,1,1)	4075,66
	SARIMA(4,0,18)(0,1,1)	4075,32
	SARIMA(10,0,9)(0,1,0)	4283,52
Los Angeles	SARIMA(20,0,14)(1,1,0)	3796,42
	SARIMA(13,0,14)(1,1,1)	3735,64
	SARIMA(13,0,14)(0,1,1)	3733,59
	SARIMA(13,0,20)(0,1,0)	3884,346

$$(1-\phi_1 B^1-\phi_3 B^3-\phi_6 B^6)x_t=(1-\Theta_1 B^s)(1-\theta_1 B^1-\theta_3 B^3-\theta_3 B^3-\theta_6 B^6)\varepsilon_t \qquad \text{(B.5)}$$

Thus, only the latter is maintained as valid models to be used in the forecasting of the solar radiation and wind speed for a duration of two years.

B.1.2/ Forecasting evaluation

Now coming back to the objective to predict the future meteorological data with the valid SARIMA$(p,d,q)(0,1,1)$ model obtained in the section before, the results are shown in Figure B.3. Moreover, to investigate the model sufficiency, we summarize the useful statistics about the forecasting results in Table B.3 by computing the mean absolute percentage error of all the tested weeks.

The forecasting model is applied on four years where it compares with the last two years 2011 and 2012 and continue the forecasts till 2014 weekly.

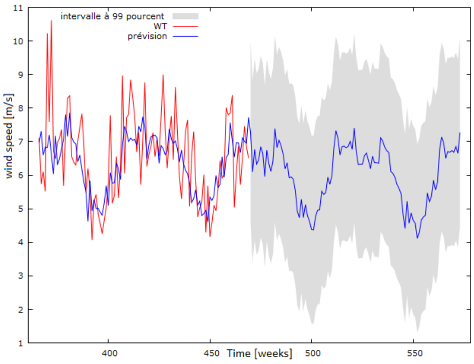
Based on the solar radiation prevision in Figures B.3b and B.3d, the forecasting curves in blue fits to the real data of the years 2011 and 2012 and follow the seasonality. Moreover, with a MAPE equal to 7 and 15.60 for the two sites Los Angeles and Chicago respectively, the SARIMA model is quite validated and accurate.

Table B.3: The mean absolute percentage error of the used methods

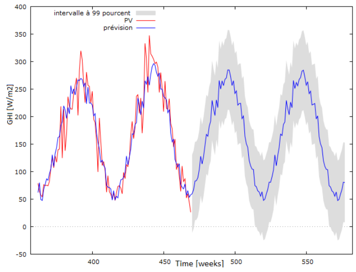
Data	Locations	Methode	MAPE
Solar radiation	LA	SARIMA(4,0,18)(0,1,1)	7,03
	Chicago	SARIMA(11,0,14)(0,1,1)	15,60
Wind Speed	LA	SARIMA(6,0,6)(0,1,1)	29,83
	Chicago	SARIMA(13,0,14)(0,1,1)	12,54

Figures B.3a and B.3c show the wind speed forecasting starting from 2011 till 2014. The forecasting curves in blue is following the trend of the real data of the years 2011 and 2012. Nevertheless, the 99% confidence interval is quite large and indicates high variability. Thus, these rates should be interpreted with the noise variance estimated. The SARIMA model applied on the wind speed in these two cities, Los Angeles and Chicago, is not as precise as the solar forecasting.

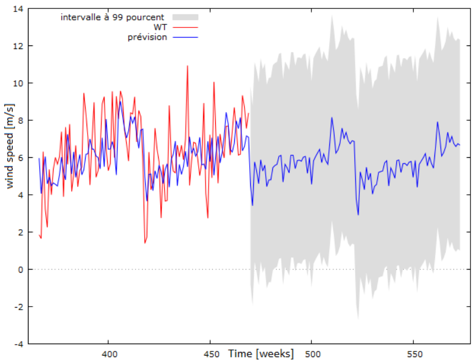
Finally, recall that all interpretations and conclusions presented in this paper are based on



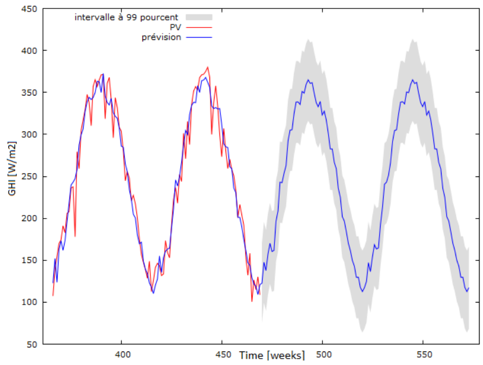
(a) Wind speed in Chicago



(b) Solar radiation in Chicago



(c) Wind speed in Los Angeles



(d) solar radiation in Los Angeles

Figure B.3: Forcasting results of the wind speed and solar radiation for Chicago and Los Angeles

data available for the specific areas. Results can vary from an area to another depending on different parameters (pressure, temperature, etc.).

B.2/ Synthesis

This appendix presents a comparison among four distinct solar radiation and wind speed generation forecasting models. It is shown that in general, SARIMA model is quite good in the forecasting of the solar radiation during years and fits very well the data because of their seasonal distribution. We also pointed out that the performance of the used model in forecasting the solar during the years is more precise than the ones for the wind speed which degrades noticeably for long term previsions. Moreover, the construction of the wind model depend strongly on the geographical area, which is less important for the sun It is hence important to predict wind speed variation as precisely as possible. This shows the interest to consider other models or characteristics such as Markov Switching ARMA [44] to improve the precision of the results in order to get an optimal sizing and management for the hybrid renewable energy system supplying a data center power demand.