Diego A.
Tejada
Arango

Escuela Técnica Superior de Ingeniería (ICAI)

CO-OPTIMIZATION OF ENERGY STORAGE TECHNOLOGIES IN TACTICAL AND STRATEGIC PLANNING MODELS

Autor: Diego Alejandro Tejada Arango
Director: Efraim Centeno Hernáez
Codirector: Sonja Wogrin

CO-OPTIMIZATION OF ENERGY STORAGE TECHNOLOGIES IN TACTICAL

MADRID | Julio 2019

CONSTANCIA REGISTRAL DEL TRIBUNAL DEL ACTO DE LA DEFENSA DE TESIS DOCTORAL

TÍTULO: Co-optimization of Energy Storage Technologies in Tactical and Strategic Planning Models

AUTOR: MSc. Diego Alejandro Tejada Arango

DIRECTOR: Dr. Efraim Centeno Hernáez

CODIRECTOR: Dra. Sonja Wogrin

TUTOR-PONENTE:

DEPARTAMENTO: Instituto de Investigación Tecnológica (IIT)

FACULTAD: Escuela Técnica Superior de Ingeniería (ICAI)

Miembros del Tribunal Calificador:

PRESIDENTE: Dr. Benjamin Hobbs	**Firma:**
VOCAL: Dr. Sara Lumbreras Sancho	**Firma:**
VOCAL: Dr. Afzal S. Siddiqui	**Firma:**
VOCAL: Dr. Germán Morales-España	**Firma:**
SECRETARIO: Dr. Andrés Ramos Galán	**Firma:**

Fecha de lectura: 3 de julio de 2019

Calificación:

"As you set out for Ithaka
hope the voyage is a long one,
full of adventure, full of discovery.
Laistrygonians and Cyclops,
angry Poseidon—don't be afraid of them:
you'll never find things like that on your way
as long as you keep your thoughts raised high,
as long as a rare excitement
stirs your spirit and your body.
Laistrygonians and Cyclops,
wild Poseidon—you won't encounter them
unless you bring them along inside your soul,
unless your soul sets them up in front of you.

Hope the voyage is a long one.
May there be many a summer morning when,
with what pleasure, what joy,
you come into harbors seen for the first time;
may you stop at Phoenician trading stations
to buy fine things,
mother of pearl and coral, amber and ebony,
sensual perfume of every kind—
as many sensual perfumes as you can;
and may you visit many Egyptian cities
to gather stores of knowledge from their scholars.

Keep Ithaka always in your mind.
Arriving there is what you are destined for.
But do not hurry the journey at all.
Better if it lasts for years,
so you are old by the time you reach the island,
wealthy with all you have gained on the way,
not expecting Ithaka to make you rich.

Ithaka gave you the marvelous journey.
Without her you would not have set out.
She has nothing left to give you now.

And if you find her poor, Ithaka won't have fooled you.
Wise as you will have become, so full of experience,
you will have understood by then what these Ithakas mean."

C.P. Cavafy, Collected Poems. Translated by Edmund Keeley and Philip Sherrard.
Edited by George Savidis. Revised Edition. Princeton University Press, 1992

Acknowledgments

Tal como el poema de Cavafy sobre Ithaka, esta tesis ha representado para mí un largo viaje lleno de experiencias, conocimiento, alegrías y dificultades que lo han hecho único e irrepetible. Para llegar hasta aquí he tenido la fortuna de contar a mi lado con mi familia, maestros y amigos, tanto los de toda la vida como los nuevos que he encontrado en este camino. A todos ellos gracias, ya que este espacio se queda un poco corto para mencionarlos a todos de forma explícitamente.

En mi familia quiero agradecer especialmente a Nubia Toro (mi abuela) y mi mamá, que son las personas que más me han enseñado con su ejemplo y apoyado en todas mis decisiones. También a mis tíos Gustavo y Carlos, que siempre han sido personas a las que he admirado mucho y a las que siempre he tenido de referencia para llegar hasta acá.

Este viaje tampoco se podría haber completado sin el apoyo, guía y enseñanzas de mis supervisores, Efraim y Sonja. La experiencia y el coaching de Efraim, así como las ideas y la pasión por la investigación de Sonja, han sido claves de este éxito personal y profesional. A ellos dos mil gracias por haber generado un ambiente ideal para desarrollar esta tesis.

Debo agradecer también a mis amigos de toda la vida y sus familias, Pablo, Juliana, Conavi, Fabián, Leonardo, Juan Esteban, Mauricio, Esteban y Jaime. Ellos han estado presentes apoyándome en muchas etapas de mi vida, y esta no ha sido la excepción. También quiero agradecer, por sus contribuciones directas o indirectas a esta tesis, a los nuevos amigos que he conseguido a lo largo de este viaje, entre ellos quiero destacar a Aurora, Eva, Maya, Dean, Rodrigo de Marcos, Isaac, Cristian, Carlos Mario, Germán y todos miembros del IIT, especialmente a los de la segunda planta.

Gracias a Diana por ser la chispa que encendió la luz hacia este camino.

Finalmente, evocando a Cerati, quiero decir que no solo no hubiera sido posible haber realizado este viaje sin ustedes, sino con toda la gente que estuvo a mi alrededor desde el comienzo, algunos siguen hasta hoy ¡gracias totales!

Summary

Nowadays, power systems are transitioning to an increasing penetration of vast low-cost wind and solar generation in order to achieve the greenhouse-gas-emission reduction targets in the electricity sector, which will require system flexibility for balancing requirements to maintain system performance. The current technologies have limited technical capabilities to provide this flexibility, and new alternatives are required. In this context, energy storage is one of the most promising options that can deliver technical and economic benefits. However, modeling energy storage systems represents a challenge because they have a wide range of technologies from pumped hydro units to batteries, each one with different characteristics that make them more suitable either short- (e.g., hours) or long-term (e.g., months) applications.

This thesis proposes optimization models that improve current operational and investment planning tools by a better consideration of short- and long-term operational decisions for different grid-level energy storage technologies that impact tactical and strategic planning in power systems. This thesis then tackles the energy storage operation and investment problem in the following aspects:

- *Representation of Energy Storage Operation*: we propose improvements in current decision support models to deal with short-term storage such as batteries and seasonal storage at the same time, including network-constrained analysis. In addition, it determines the main drawbacks of the traditional modeling approaches using an hourly unit commitment model as a benchmark for the comparison of the current and proposed models.
- *Co-optimization of Energy Storage Technologies in hydrothermal dispatch models*: we assess the impact of short-term energy storage decisions on the opportunity cost of long-term storage through the proposal of a new optimization model for hydrothermal coordination in which hourly opportunity costs or short-term signals are co-optimized with seasonal storage.
- *Investment Decision Models for Energy Storage*: we formulate and test the main modeling approaches to evaluate energy-storage-systems investment in power systems with high penetration of renewable energy sources. Moreover, we analyze the influence of transmission constraints, losses, and increased renewable energy penetration on planning energy-storage-systems allocation and investment.
- *Investment Decision Models for Energy Storage using Power-based Unit Commitment*: we improve current investment models by correctly modeling power system flexibility requirements that lever different energy storage investment. Moreover, we compare energy-based and power-based unit commitment models and analyze the main advantages and disadvantages for the energy storage investment decisions.

The proposed models can support energy storage owners, investors, system operators, planning entities, and regulatory authorities in their decisions regarding energy storage in the future context of high share of variable renewable energy sources.

Dissertation

This doctoral thesis analyzes the energy storage systems (ESS) modeling in power system tactical and strategic planning (also known as medium- and long-term planning) with emphasis on the co-optimization of different ESS technologies considering two main aspects: operation and investment. This document is based on the work of 7 articles (4 in journals with JCR, 1 in journals without JCR, 2 under review), which are included at the end of this document (labelled Article I–VII) and listed as follows. Further details of the thesis structure and roadmap are given in Section 1.4.

Journal Articles

Article I. Tejada-Arango, D. A., Wogrin, S., & Centeno, E. (2018). Representation of Storage Operations in Network-Constrained Optimization Models for Medium- and Long-Term Operation. *IEEE Transactions on Power Systems*, 33(1), 386–396.

Article II. Tejada-Arango, D. A., Domeshek, M., Wogrin, S., & Centeno, E. (2018). Enhanced Representative Days and System States Modeling for Energy Storage Investment Analysis. *IEEE Transactions on Power Systems*, 33(6), 6534–6544.

Article III. Yacar, D., Tejada-Arango, D., & Wogrin, S. (2018). Storage Allocation and Investment Optimization for Transmission Constrained Networks Considering Losses and High Renewable Penetration. *IET Renewable Power Generation,* 12(16), 1949-1956.

Article IV. Tejada-Arango, D. A., Siddiqui, A. S., Wogrin, S., & Centeno, E. (2019). A Review of Energy Storage System Legislation in the US and the European Union. *Current Sustainable/Renewable Energy Reports,* 6(1), 22-28.

Article V. G. Morales-Espana and D. A. Tejada-Arango, "Modelling the Hidden Flexibility of Clustered Unit Commitment," *IEEE Transactions on Power Systems*, pp. 1–1, 2019.

Under Review

Article VI. D.A. Tejada-Arango, S. Wogrin, A. Siddiqui, E. Centeno, Short-term storage signals in hydrothermal dispatch models using a linked representative periods approach, October 2018. (currently under review in Energy - The International Journal).

Article VII. D.A. Tejada-Arango, G. Morales-España, S. Wogrin, E. Centeno, Power-Based Generation Expansion Planning for Flexibility Requirements, February 2019. (currently under review in IEEE Transactions on Power Systems).

In addition, during my process as research assistant at the IIT, I have published three other (JCR) journal papers [1]–[3].

Contents

List of Figures

List of Tables

Acronyms

AE	Accumulated Error
BESS	Battery Energy Storage Systems
CAES	Compressed Air Energy Storage
CUC	Clustered Unit Commitment
DAM	Day-Ahead Markets
DSO	Distribution System Operator
ENTSO-E	European Network of Transmission System Operators for Electricity
EPR	Energy to Power Ratio
ESS	Energy Storage Systems
EU	European Union
FERC	Federal Energy Regulatory Commission
FES	Flywheel Energy Storage
GEP	Generation Expansion Planning
IEA	International Energy Agency
ISO	Independent System Operator
LDC	Load Duration Curve
LRP	Linked Representative Periods
LTESS	Long-Term Energy Storage Systems
MLTOP	Medium- and Long-Term Operational Planning
NRMSE	Normalized Root Mean Squared Error
OPF	Optimal Power Flow
PHS	Pump Hydro Storage
RP	Representative Periods
RP-TM&CI	Representative Periods with Transition Matrix and Cluster Indices
RTO	Regional Transmission Operator
SED	Squared Euclidean Distance
SS	System States
SS-RFM	System States Reduced Frequency Matrix
STESS	Short-Term Energy Storage Systems
TSO	Transmission System Operator
UC	Unit Commitment
US	United States
VRES	Variable Renewable Energy Sources

1 INTRODUCTION

1.1 Motivation and Context

The effects of global warming and climate change have mobilized several countries around the world to mitigate and reduce the greenhouse gas emissions, making the transition into a decarbonized power system one of the main challenges of the 21st century in the electric power industry. This has led to the integration of renewable energy sources such as solar and wind power. In the last decade, the installation of variable renewable energy sources (VRES) has continuously increased in almost all electric power systems around the world. This is mainly due to regulatory incentives as well as reductions on renewable technology manufacturing costs. In addition, according to the International Energy Agency (IEA) renewables account for 80% of new capacity in the European Union, and by 2040, they will represent two-thirds of global investment in power plants [4]. The main advantages of renewable energy sources are their low variable cost and clean energy production, i.e., without causing CO2 emissions, unlike traditional thermal generation sources whose fuel is coal or natural gas. However, high penetration of renewables in the electricity sector also has important negative side effects [5], such as:

- *Non-controllable variability*: the total output in VRES depends on the renewable source availability (e.g., wind flow or sunlight). Thus, VRES cannot produce a constant output within the short-term planning horizons (e.g., day-ahead unit commitment). Thus, system operators face changes on the VRES planned production in real-time operation that usually are controlled using other generation resources.
- *Partial unpredictability*: Although forecasting renewable production has improved in the last years [6], climate change and large-scale ocean-atmosphere climate interactions, such as El Niño y La Niña [7], still lead to an uncertainty the VRES production.
- *Locational dependence*: VRES are often located in remote locations (especially onshore or offshore wind farms), far from load centers. Therefore, current transmission infrastructure will be more congested while new transmission investments are developed.

This situation brings additional problems to the operation of the system, such as flexibility reduction because of this non-dispatchable characteristic, the intermittent output of this renewable energy, and the increased level of uncertainty in their production. Energy storage systems (e.g., pumped hydro storage, compressed air energy storage, or batteries) are one of the options to increase the flexibility of power systems because they can store energy at off-peak times and discharge it at peak times (i.e., time shifting) [8]. Therefore, energy storage systems represent a suitable candidate to soften the transition to decarbonization of power systems and, therefore, achieving a "clean electrical power industry". In fact, thermal systems, renewable technologies, and energy storage technologies are already co-existing in several electric power systems. Nevertheless, new energy storage technologies, e.g., batteries, rep-

resent a small part of the total installed generation capacity due to their current investment cost. However, these investment costs are expected to decrease in the forthcoming years, as well as having different stream of revenues due to multiple services of energy storage in the power system (e.g., reserves, frequency control, capacity mechanisms) due to the large-scale penetration of renewable energy sources. This situation will boost the integration of energy storage system (e.g., batteries or equivalent storage technologies) in order to achieve the full decarbonization of power systems. Leading to a need of adaptation of current planning models in order to consider operational and investment decisions in this new context.

This thesis aims to improve current operation and investment optimization tools by properly modeling different types of energy storage systems technologies. Short-term characteristics of energy storage systems can provide solutions to diurnal generation cycles of renewable energies that do not match load cycles. Advantages of renewable energy integration due to energy storage systems have been analyzed in [5], [9]. This shows that energy storage could improve operational flexibility in the decarbonization of the power sector. However, short-term energy storage (e.g., battery energy storage systems) has not been analyzed and co-optimized with seasonal storage, which represents a challenge to properly consider the chronological constraints of both types of storage at the same time. In addition, short-term energy storage, e.g., batteries, could provide multiple services in power systems such as: load shifting (a.k.a. energy arbitrage), renewable support, reserve markets, etc.; and consequently, new ways to consider these multiple services in the planning tools or decision support models are needed. As part of the state-of-the-art research in this thesis, in Article IV [10] we have assessed recent regulatory proposals in the US and the EU in order to understand their implications for ESS providing multiple services. Together with the other publications in this thesis, full content of Article IV [10] is included at the end of this document.

Among the different power system planning models, there are short-term models with high resolution times such as unit commitment models, with information pertaining to every hour, half hour, or 10 minutes; and long-term models such as investment models that ignore small time-scale changes to make the calculations in a reasonable amount of time. The introduction of VRES into the energy system, however, makes it necessary to include more short-term dynamics in long-term models [11]. Realistically modeling energy storage requires the preservation of chronological information, because the amount of stored energy available at any given moment depends on the amount of energy stored in all previous time periods [12]. However, most of the current models and tools for tactical and strategic planning do not fully consider these chronological constraints:

- *Tactical planning*: The hydrothermal dispatch has been one of the most studied problems in tactical planning [13]–[16] due to the hydro inflows uncertainty for seasonal storage. The main tools that are available to perform a hydrothermal dispatch analysis are: SDDP developed by PSR [17], PLEXOS® Integrated Energy Model developed by Energy Exemplar [18], ProdRisk developed by SINTEF [19], and StarNet Model developed by IIT

[20]. For medium- or long-term studies, these tools use a load levels approach with monthly or weekly stages. Moreover, no relevant effort has been performed in the last years to improve the short-term operational decisions on hydro-thermal dispatch models. Therefore, there is a lack of studies and analyses of the economic impact of short-term storage operational decisions on seasonal storage opportunity costs in hydrothermal dispatch problems such as Moreno et al. explain in [21].

- *Strategic planning*: Investment decisions are one of the most important decisions in strategic planning [22]–[24]. Investment models that incorporate information at both time scales include the TIMES modeling framework [25], the Regional Energy Deployment System (ReEDS) framework [26], Resource Planning Model (RPM) [27], and COMPETES [28]. These models have multi-year investment decisions as well as *time slices* within each year that represent a wide variety of possible demand and VRES production levels. Although some models have endeavored to incorporate energy storage investment and operational decisions, they do not preserve chronological information and thus do not fully model storage evolution [29], [30].

Finally, current tactical and strategic planning models use energy-based formulations. However, recent studies [31], [32] have shown that energy-based formulations overestimate the actual flexibility of the system. Instead, power-based UC models overcome these problems by correctly modeling operating reserves and ramping constraints. As mentioned before, energy storage systems are important to increase the power system flexibility, therefore, a proper modeling of energy storage systems in the power-based formulations is needed. However, modeling of energy storage systems in the power-based formulation has not been analyzed nor considered so far.

1.2 Objectives

1.2.1 Main Objective

The main objective of this thesis is to improve current operational and investment planning tools for a better consideration of short- and long-term decisions for different grid-level energy storage technologies that affects tactical and strategic planning in power systems.

1.2.2 Specific Objectives

The main objective can be broken down into the following specific objectives:

Objective 1. *Representation of Energy Storage Operation*: To propose improvements in current decision support models to deal with short-term storage such as BESS and seasonal storage at the same time, including network-constrained analysis. In addition, determine the main drawbacks of the traditional modeling approaches using an hourly unit commitment model as a benchmark for the comparison of the current and proposed models.

Objective 2. *Co-optimization of Energy Storage Technologies in hydrothermal Dispatch Models*: To assess the impact of short-term energy storage decisions on the opportunity cost of long-term storage through the proposal of a new optimization model for hydrothermal coordination (i.e., considering the uncertainty in the natural hydro inflows) in which hourly opportunity costs or short-term signals are co-optimized with seasonal storage.

Objective 3. *Investment Decision Models for Energy Storage*: To formulate and test the main modeling approaches to evaluate energy-storage-systems investments in power systems with high penetration of renewable energy sources. Furthermore, to analyze the influence of transmission constraints, losses, and increased renewable energy penetration on planning energy-storage-systems allocation and investment.

Objective 4. *Investment Decision Models for Energy Storage using Power-based Unit Commitment*: To improve current investment models by correctly modeling power system flexibility requirements (mainly due to VRES production) that lever different ESS technologies investment. Moreover, compare energy-based and power-based unit commitment models and analyze the main advantages and disadvantages for the ESS investment decisions.

1.3 Thesis Structure

This thesis focuses on the co-optimization of energy storage technologies for medium- and long-term planning tools, especially on modeling of operational and investment decisions, see Figure 1-1. To achieve this, it is necessary to properly consider the chronological constraints of ESS while maintaining a tractable solution time.

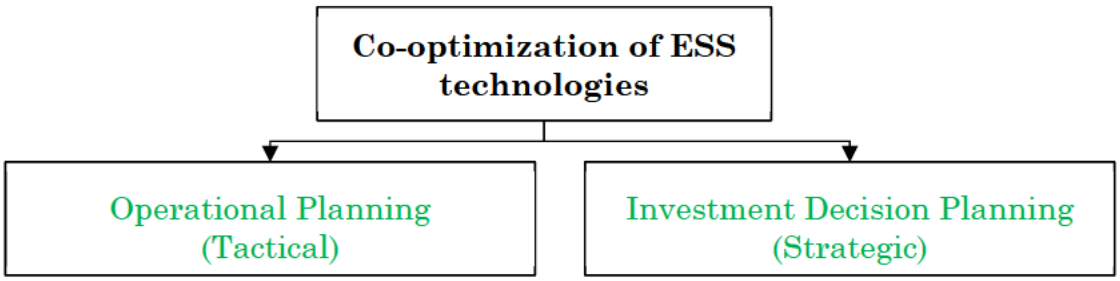


Figure 1-1. Thesis Structure

First, for the ESS *operational planning*, we analyze the main available methods to reduce the computational time in medium- and long-term models that consider the ESS constraints. We review the main limitations in these methods and propose enhanced versions to overcome these limitations. In particular, we focus on the representation of the short-term operational decisions by using a Unit Commitment model as a benchmark. Therefore, short-term operational decisions are detailed in our proposed models, allowing to co-optimize the operational decisions of both intra-day and inter-day storage. In addition, most of the current reduction techniques for time representation rely on clustering techniques and yet the impact of the transmission network on

these clustering techniques has not been adequately analyzed. We analyze this impact to draw some conclusions and give some recommendations about this matter. We finish the analysis of operational planning in ESS by extending the proposed models to a stochastic hydrothermal dispatch model to determine the impact of intra-day storage on the seasonal storage (i.e., inter-day storage). The storage and water value are analyzed in the proposed models in order to obtain the economic signals of energy storage in both types of ESS.

Second, for the ESS *investment decision planning*, once we have ensured a good representation of operational decisions for the intra-day and inter-day storage, it is possible to analyze the investment decisions for different ESS technologies in a more accurate manner. We test the proposed models including investment decisions for different ESS technologies. In addition, we study the impact of two important aspects in the ESS investment: 1) transmission congestion and losses, 2) Power-based UC for flexibility representation in power systems. Regarding transmission congestion and losses, they are traditionally neglected in investment decision planning, however, we discuss the main implications for ESS investment once a more detailed transmission network is considered. Power-based UC models are getting more and more relevance for their accuracy in the representation of power system flexibility requirements. Since, ESS are one of the options to increase the flexibility in power systems, we investigate the investment decisions in ESS with a more detailed representation of these requirements considering different types of ESS technologies. Therefore, we provide a wide range of aspects that impact the ESS investment decisions in different types of technologies.

Finally, this thesis proposes computational efficient optimization models to determine the operation and investment decisions of different types of ESS technologies at the same time. These models can support ESS owners, investors, ISOs, planning entities, and regulatory authorities in their decision-making process regarding ESS in the future context of high share of VRES.

1.4 Thesis Outline

This section presents the roadmap of the thesis. This roadmap or outline is based on a partition into two parts, each of which covers two specific objectives. In addition, each objective is supported by one or more scientific contributions (i.e., articles or papers). We associate the papers with the different objectives, nevertheless, some of the following papers relate to more than one objective, as shown in Figure 1-2.

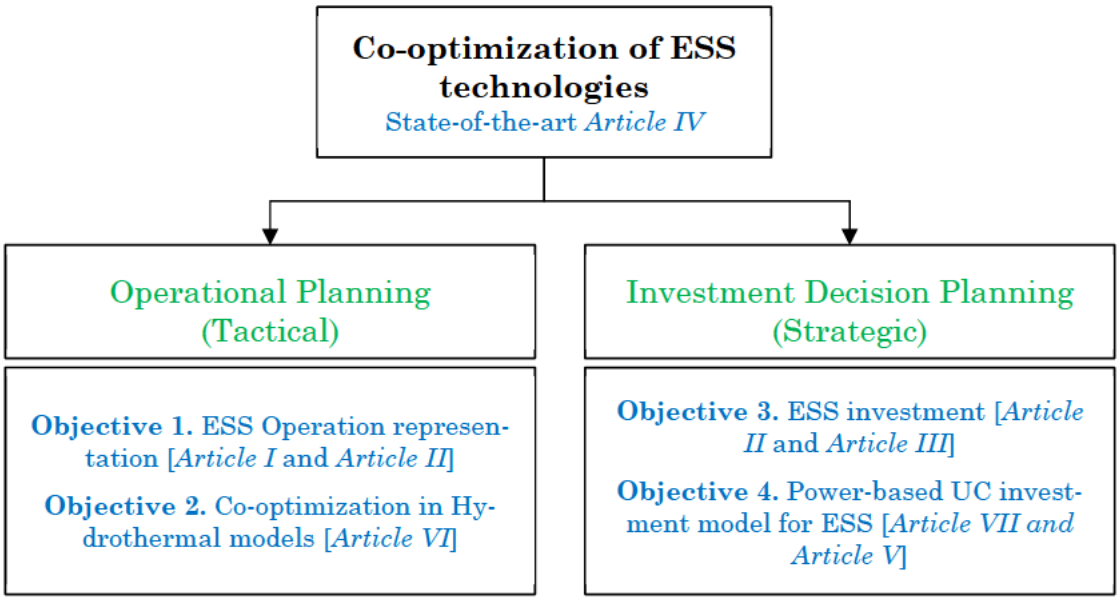


Figure 1-2. Relationship among structure, objectives, and contributions

Article I. This paper proposes a model to carry out analysis of storage facilities operation including a transmission network. The model represents short-term storage operation in an approximated way that reduces computational requirements, which makes it suitable for operational planning in power systems with a high level of renewable energy penetration. In the proposed model, we cluster hourly data using the so-called system-states framework developed in previous work. Within this framework, non-consecutive similar time periods are grouped, while chronological information is represented by a transition matrix among states. We extend the system-state framework from a single-bus system to a transmission network. We define and analyze two alternative sets of representative variables for clustering hours to obtain system states when the transmission network is considered. This extension of the system states framework allows us to evaluate the impact of transmission congestions in medium- and long-term planning models in a reasonable computation time. A case study shows that the proposed model is 235 times faster than an hourly approach, used as benchmark, whereas the overall system cost is approximated with less than 2% error. The overall charging/discharging trends are similar enough to those of the hourly model, being hydro storage better approximated than fast-ramping batteries. Besides, for the analyzed case study, it is shown how congestion in the transmission net-work in fact improves the accuracy of the proposed approach.

Article II. This paper analyzes different models for evaluating investments in ESS in power systems with high penetration of VRES. First of all, two methodologies proposed in the literature are extended to consider ESS investment: a unit commitment model that uses the *System States* (SS) method of representing time; and another one that uses a *Representative Periods* (RP) method. Besides, this paper proposes two new models that improve the previous ones without a significant increase of computation time. The enhanced models have been called the *System States Reduced Frequency Matrix* (SS-RFM) model which addresses

short-term energy storage more approximately than the SS method to reduce the number of constraints in the problem, and the *Representative Periods with Transition Matrix and Cluster Indices* (RP-TM&CI) model which guarantees some continuity between representative periods, e.g. days, and introduces long-term storage into a model originally designed only for the short term. All these models are compared using an hourly unit commitment model as benchmark. While both system state models provide an excellent representation of long-term storage, their representation of short-term storage is frequently unrealistic. The RP-TM&CI model, on the other hand, succeeds in approximating both short- and long-term storage, which leads to almost 10 times lower error in storage investment results in comparison to the other models analyzed.

Article III. This paper investigates the effects of transmission losses, constraints and increased renewable energy penetration on planning energy storage allocation and investment. By modifying a DC Optimal Power Flow model using a linearized approximation for ohmic losses we were able to understand which network characteristic or inhibitor drives the most change in expanding utility scale storage. Four different storage technologies were explored: Compressed Air Energy Storage, Pumped Hydro Storage, Lithium-Ion Battery and Fly Wheel. Each had different charging, capacity and cost characteristics. The results of the storage allocation trials revealed that network congestion was a more influential network inhibitor than were line losses. Losses only had substantial effects on a free-flowing network but produced marginal changes in allocation in congested ones. The conclusion of the investment trials revealed two things: 1) Storage investment is not significantly affected by transmission constraints so long as renewable generation stays constant and relatively low; 2) More flexible technologies like Flywheels are favored at lower volumes of renewable penetration for their load balancing capabilities while cheaper technologies are best as the volume of renewable power generated increases.

Article IV. This paper focuses on the current possibilities for energy storage systems (ESS) to participate in different power system services. ESS can provide multiple services such as spinning reserve, deferral upgrades, and energy management. However, this versatility of ESS poses a challenge for regulators in designing markets where ESS have prominent roles. We assess recent regulatory proposals in the US and the EU in order to understand their implications for ESS. These proposals attempt to improve the current rules for efficient ESS deployment. Nevertheless, they have different approaches to the same problem. We discuss these differences in an attempt to shed light on the regulatory debate about ESS ownership and market design. The successful integration of ESS will depend on proper incentives to provide multiple services without hampering the current market structure. New asset definitions could help to define the roles of ESS as either a generation or a transmission asset.

Article V. This paper proposes a Clustered Unit Commitment (CUC) formulation to accurately model flexibility requirements such as ramping,

reserve, and startup/shutdown constraints. The classic CUC intrinsically and hiddenly overestimates the individual unit's flexibility, thus being unable to replicate the result of the individual UC. This paper presents a set of constraints to correctly represent the units' hidden flexibility within the cluster. Different case studies show that the proposed CUC replicates the results of the individual UC while solving significantly faster. Therefore, the proposed CUC correctly represents the individual unit's flexibility within the cluster and could be used in large-scale planning models without significantly increasing their computational burden.

Article VI. Short-term energy storage systems (STESS), e.g., batteries, are becoming one promising option to deal with flexibility requirements in power systems due to the accommodation of renewable energy sources. Previous work using medium- and long-term planning tools has modeled the interaction between STESS and seasonal storage (e.g., hydro reservoirs). Despite these developments, opportunity costs considering the impact of STESS signals in stochastic modeling have not been analyzed. This paper proposes a new formulation to include STESS operational decisions in a stochastic hydrothermal dispatch model, which is based on Linked Representative Periods (LRP) approach that allows an analysis of both short- and long-term storage at the same time. This proposal models operating decisions of STESS with errors between 5% to 10%, while the classic Load Duration Curve (LDC) approach fails by an error greater than 100%. Moreover, the LDC model cannot determine opportunity costs on an hourly basis and underestimates the water value by 6% to 24% for seasonal hydro reservoirs. On the other hand, the proposed LRP model produces an error on the water value lower than 3% and can determine hourly opportunity costs for STESS using dual variables from both intra- and inter-period storage balance equations. Therefore, hourly opportunity costs in the LRP model successfully internalize long-term signals due to seasonality in hydro reservoirs.

Article VII. Flexibility requirements are becoming more relevant in planning process due to the integration of variable Renewable Energy Sources (VRES). In order to consider these requirements Generation Expansion Planning (GEP) models have recently incorporated Unit Commitment (UC) constraints, using energy-based formulations. However, recent studies have shown that energy-based UC formulations overestimate the actual flexibility of the system. Instead, power-based UC models overcome these problems by correctly modeling operating reserves and ramping constraints. This paper proposes a power-based GEP-UC model that improves the existing models. The proposed model optimizes investment decisions on VRES, Energy Storage Systems (ESS), and thermal technologies. In addition, it includes transmission network constraints, real-time flexibility requirements, and the flexibility provided by ESS. The results show that power-based model uses the installed investments more effectively than the energy-based models because it is more accurate in the representation of flexibility re-

quirements. For instance, the power-based model obtains less investment (6-12%) and yet it uses more efficiently this investment because operating cost is also lower (2-8%) in a real-time validation. We also propose a semi-relaxed power-based GEP-UC model, which is at least 10 times faster than its full-integer version and without significantly losing accuracy in the results (less than 0.2% error).

Finally, the structure of this document is described as well as a summary of all models that have been developed in this thesis.

Chapter 2. *Background*. This chapter shows the basic background of the thesis research topics. It shows an overview of the different available ESS technologies, which is followed by a discussion on the common approaches to deal with the chronological constraints. It analyzes the main advantages and disadvantages of the common approaches in order to determine the main gaps in the current state-of-the-art in this topic. In addition, it shows a general overview of the unit commitment problem, which is used as the benchmark formulation to determine the operational decisions in the proposed models of this thesis. Finally, it discusses the main challenges in the representation of ESS for flexibility requirements in unit commitment formulations.

Chapter 3. *ESS operation modeling*. This chapter aims at developing Objective 1 of this thesis and it is based on the analysis and results in Article I and Article II. Optimization models for ESS operation in medium- and long-term planning are studied in two main parts. The first part analyzes the impact of transmission network in the clustering methods used to reduce the temporal information while conserving chronological information for the ESS. The second part compares the system states and representative periods methods when short-term energy storage and seasonal storage are included in the optimization models. In addition, enhanced versions of each method are proposed in order to overcome the main drawbacks in their former methods.

Chapter 4. *Hydrothermal dispatch using linked representative periods*. This chapter aims at developing the Objective 2 of this thesis and it is based on the analysis and results in Article VI. This chapter proposes a new model to include short-term ESS operational decisions in a stochastic hydrothermal dispatch model, which is based on the proposed enhanced version of the representative periods approach that allows an analysis of both short- and long-term storage at the same time. This proposal models operating decisions of short-term ESS, while the classic Load Duration Curve (LDC) approach fails on this representation. Moreover, the LDC model cannot determine opportunity costs on an hourly basis and underestimates the water value for seasonal hydro reservoirs. The proposed model can determine hourly opportunity costs for short-term ESS using dual variables from both intra- and inter-period storage balance equations. Therefore, hourly opportunity costs in the proposed model successfully internalize long-term signals due to seasonality in hydro reservoirs.

Chapter 5. *ESS investment modeling*. This chapter aims at developing the Objective 3 of this thesis and it is based on the analysis and results in

Article II and Article III. Therefore, it analyzes optimization models for ESS investment in medium- and long-term planning in two main parts. The first part investigates the effects of transmission losses, constraints and increased renewable energy penetration on planning ESS allocation and investment. By modifying a DC Optimal Power Flow model using a linearized approximation for ohmic losses we were able to understand which network characteristic or inhibitor drives the most change in expanding utility scale storage. The second part compares the system-states type and representative-periods type methods when short-term energy storage and seasonal storage are included in the optimization models for ESS investment decisions. Finally, it analyzes the close relationship between renewable curtailment and ESS investment in both type of models.

Chapter 6. *Power-Based Model for Flexibility Requirements.* This chapter aims at developing the Objective 4 of this thesis and it is based on the analysis and results in Article V and Article VII. Therefore, it proposes a power-based GEP-UC model that improves the classic energy-based models by representing more accurately the flexibility requirements of power systems (i.e., reserve decisions and ramping constraints), as well as a real-time validation stage (e.g., 5-min simulation) in order to evaluate the quality of investment and operational decisions obtained with the model. Moreover, it also proposes a semi-relaxed version of the power-based GEP-UC model, which aims at reducing the computational burden without losing accuracy in the results.

Chapter 7. The last chapter of this thesis draws the conclusions and remarks possible future works on this topic.

Table 1-1. Summary of optimization models in this thesis

Model[1]	Type of Planning Section		Temporal representation for one target year[2]	ESS technologies		Unit Commitment	Case study[3]				Annex with equations
	Tactical	Strategic[4]		Short Term ESS	Long Term ESS		IEEE 14-bus	IEEE 118-bus	Stylized Spanish case	Stylized Dutch case	
HM	3.1, 3.2, 4.3	5.2 (ESS)	8760h	BESS	Hydro	Energy-based	X	-	X	-	Deterministic (A.2) Stochastic hydrothermal (B.2)
SS	3.1, 3.2, 4.3	5.2 (ESS)	96ss	BESS	Hydro	Energy-based	X	-	X	-	Deterministic (A.3)
SS-RFM	3.2, 4.3	5.2 (ESS)	Enhanced version with 96ss	BESS	Hydro	Energy-based	-	-	X	-	Deterministic (A.4)
RP	3.2, 4.3	5.2 (ESS)	18rp (24h)	BESS	Hydro	Energy-based	-	-	X	-	Deterministic (A.5)
RP-TM&CI	3.2, 4.3	5.2 (ESS)	Enhanced version with 18rp (24h)	BESS	Hydro	Energy-based	-	-	X	-	Deterministic (A.6)
LDC	4.3	-	12ll per month	BESS	Hydro	Energy-based	-	-	X	-	Stochastic hydrothermal (B.3)
LRP	4.3	-	Enhanced version with 4rp (24h) per month	BESS	Hydro	Energy-based	-	-	X	-	Stochastic hydrothermal (B.4)
AIM	-	5.1 (ESS)	1rp (24h divided in 5-min time steps)	PSH, CAES, BESS, FES	-	Energy-based	X	-	-	-	Deterministic (C.2-C.4) including transmission losses

(continued)

Model[1]	Type of Planning Section		Temporal representation for one target year[2]	ESS technologies		Unit Commitment	Case study[3]				Annex with equations
	Tactical	Strategic[4]		Short Term ESS	Long Term ESS		IEEE 14-bus	IEEE 118-bus	Stylized Spanish case	Stylized Dutch case	
EB	-	6.2 (ESS, VRES, and thermal units)	4rp (168h)	PSH, CAES, BESS	-	Energy-based	-	X	-	X	Stochastic (D.2)
EBs	-	6.2 (ESS, VRES, and thermal units)	4rp (168h)	CAES, BESS, PHS	-	Energy-based	-	X	-	X	Stochastic (D.2) including startup/shut-down trajectories
PB	-	6.2 (ESS, VRES, and thermal units)	4rp (168h)	CAES, BESS, PHS	-	Power-based	-	X	-	X	Stochastic (D.3)
SR-PB	-	6.2 (ESS, VRES, and thermal units)	4rp (168h)	CAES, BESS, PHS	-	Power-based	-	X	-	X	Stochastic (D.3) using semi-relaxed approach in 6.4

[1] The models in this table correspond to: Hourly Model (HM), System States (SS), System State with Reduced Frequency Matrix (SS-RFM) Representative Periods (RP), Representative Periods with Transition Matrix and Cluster Index (RP-TM&CI), Load Duration Curve (LDC), Linked Representative Periods (LRP), Allocation and Investment Model (AIM), Energy-based Generation Expansion Planning (EB), Energy-based Generation Expansion Planning with startup and shutdown trajectories (EBs), Power-based Generation Expansion Planning (PB), and Semi-relaxed Power-based Generation Expansion Planning (SR-PB).

[2] The temporal representation is shown for the main case study. However, more sensitivities to this value are analyzed in the corresponding section. The units in this column stand for, h: hours, ss: system states, rp: representative periods, ll: load levels.

[3] The stylized Spanish and Dutch case studies are solved as a single node considering a fixed interchange with the neighborhood countries.

[4] The information in round brackets represents the technologies in which the model can invest.

2 BACKGROUND

Contents

This chapter shows the basic background of the research topics in four main sections, each one with general concepts that will be necessary to develop the remaining chapters in this thesis. The first section starts with an overview of the different available ESS technologies, which is followed by a discussion on the common approaches to deal with the chronological constraints within the second section. We analyze the main advantages and disadvantages of the common approaches in order to determine the main gaps in the current state-of-the-art in this topic.

The third section in this chapter shows a general overview of the unit commitment problem, which is used as the benchmark formulation to determine the operational decisions in the proposed models of this thesis. In addition, we discuss the main drawbacks of energy-based formulations for unit commitment models and how power-based formulations are recently getting more attention because the overcome this drawback, especially when flexibility requirements are relevant in power systems. We also discuss the main challenges in the representation of ESS for flexibility requirements in unit commitment formulations.

The last section presents a review on ESS regulation in the US and the EU, including the main challenges for the operation and investment of ESS technologies. Finally, we discuss these challenges and shed some light on different option to face them.

2.1 Introduction

Population growth around the world, climate change, and so-called green policies are transforming our way of producing and consuming energy. For instance, green policies are demanding increasing energy production from variable renewable energy sources (VRES) [12]. Due to government support and market reforms, wind, and solar generation has been increasing over the last decade [33]. Nevertheless, integrating these vast quantities of VRES into current electric power systems leads to several technical and economic challenges. For instance, the planning and operation of power systems is more difficult to predict and manage due to the intermittent production of VRES. Furthermore, potential VRES locations are frequently geographically scattered and rarely correlated with demand profiles. These characteristics pose challenges for voltage and frequency regulation specifically and the adequacy of power systems generally [34]. As a consequence, power systems operation and planning should become more flexible and embrace new technologies that could facilitate the integration of VRES [35]. Flexibility in power systems can be attained through many different approaches such as demand-side management, VRES curtailment, intra-day markets, integration of different energy sectors (e.g., electricity, transport, heat), reinforcement of the transmission infrastructure, addition of flexible generation technologies (e.g., open cycle gas turbines), and energy storage systems (ESS) [9], [36].

ESS are often touted as potential solutions for VRES integration [5], [37]. For instance, the *Hornsdale Power Reserve Battery Energy Storage System* in Jamestown, Australia is a recent prominent case because it helped to integrate wind farms in its region [38]. This case has shown that ESS can provide multiple services to integrate VRES such as load shifting (a.k.a. energy arbitrage), reserves, and frequency control ancillary services. In addition, ESS technologies have a wide range of investment costs (i.e., per power capacity and per energy capacity), losses, maximum number of cycles, ramping capacities, and efficiency [8], [39]. This leads to potential applications in power systems such as [40], [41]:

- *Generation Services*: load shifting or energy arbitrage, balancing services, frequency response services (e.g., primary, secondary, and tertiary reserve), ramping/load following, black start, firm supply in capacity markets, and VRES curtailment reduction.
- *Transmission & Distribution Services*: System reliability improvement, congestion management, and deferral upgrades.
- *End-User Services*: power quality maintenance, demand reduction, uninterruptable power supply, and back-up power.

2.2 Energy Storage Technologies and its Representation

As it was mentioned in the previous section, ESS can provide different services in the power system that could be classified in three groups: generation, network, and end-user services. This thesis focuses on operation and investment decisions for ESS in the context of high shares of VRES, considering generation services such as load shifting or energy arbitrage, reserves, and VRES curtailment reduction.

The combination of ESS and VRES improves short-term planning for power system operation as well as mitigates the VRES production uncertainty. Moreover, at very high shares of VRES, electricity will require to be stored over different time periods (minutes, hours, days, weeks, or months) [42] in order to mitigate either short-term fluctuations or seasonal variations related to VRES production. Depending on the type of technology, ESS can provide solutions to these requirements as we discuss in the following section.

2.2.1 ESS Technologies Overview

The current state-of-the-art in ESS shows a wide range of technologies with different energy-to-power ratio (EPR), which is the relationship between energy capacity and power capacity expressed as kilowatt-hour divided by kilowatts (kWh/kW) [42]. Therefore, ESS with higher values of EPR can discharge at their rated power for several hours, while ESS with lower values of EPR can discharge at their rated power only for a few seconds or minutes, see Figure 2-1. Depending on the technology, the ESS can provide different services to power systems. On the one hand, compressed air energy storage (CAES) and pump hydro storage (PHS) have discharge times in tens of hours (e.g., weeks), with high power ratings (e.g., 1000 MW). This allows them to provide bulk power management, such as electric energy time shift (e.g., arbitrage), and electric supply capacity. On the other hand, battery energy storage systems (BESS) and flywheels energy storage (FES) have lower power ratings (per module) with shorter discharge times (e.g., minutes, or few hours). In this case, these technologies fit into other kind of power system services, such as ancillary services (spinning reserves, frequency regulation, voltage support), transmission and distribution services (upgrade deferral, congestion relief), and even load shifting for short time periods (commonly within a day). Therefore, we establish two main groups of ESS technologies which we refer throughout this thesis: long-term energy storage systems (LTESS) and short-term energy storage systems (STESS). Figure 2-1 shows a general overview for conceptual purpose, however, some ESS technologies have broader power ratings and discharge times than shown. In Figure 2-1 we could classify as LTESS the ESS that are in the *bulk power management* column (i.e., pumped hydro and CAES), while ESS in the second column (i.e., *transmission & distribution grid support load shifting*) with discharge time longer than minutes are classified as STESS. ESS that could only provide *uninterruptible power supply power quality* (i.e., first column in Figure 2-1) or with their discharge time within a few seconds are out of the scope of this thesis, since power quality phenomenon in power systems is not commonly included in tactical and strategic planning models, mainly because it has a negligible impact in the medium- and long-term operational and investment decisions.

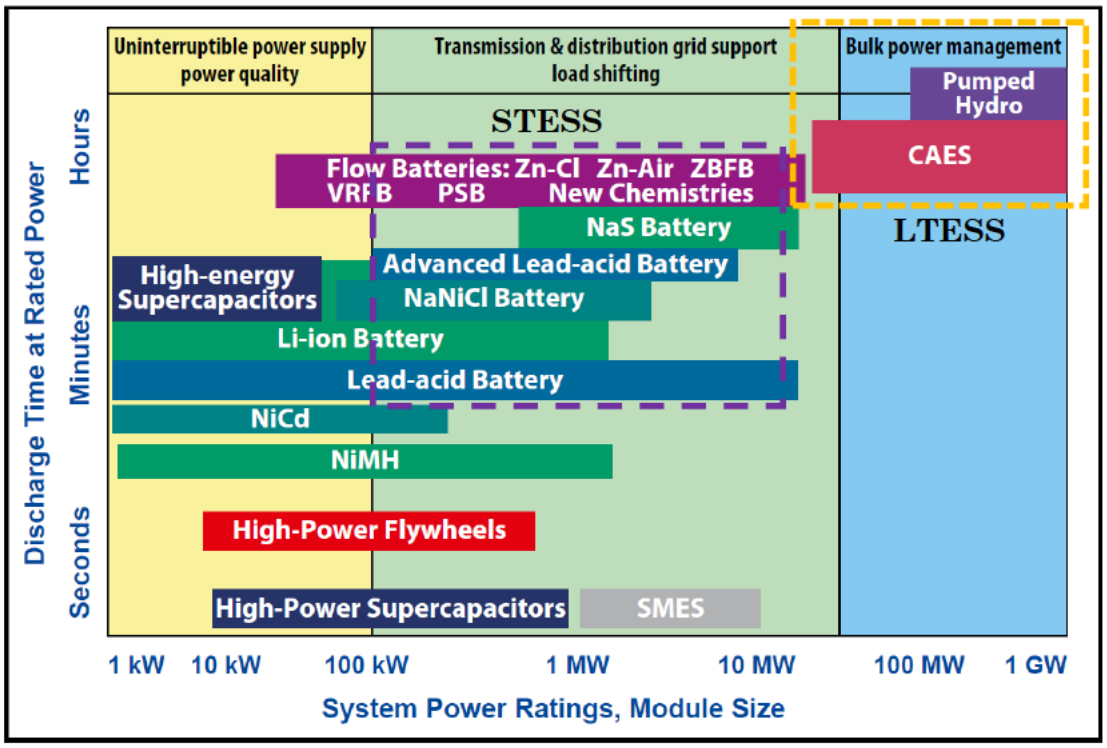


Figure 2-1. Power vs Energy Capacity for main ESS technologies [43]

2.2.2 Challenge on Chronological Constraint Representation

In the previous section, we have shown that ESS can store energy over minutes to months depending on the type of ESS technology. However, we can use a general mathematical representation for all ESS technologies:

$$\begin{aligned} &Stored\ energy\ (p) \\ &\quad = Stored\ energy\ (p-1) \\ &\quad + Charged\ energy(p) - Discharged\ energy(p) \end{aligned}$$

Therefore, the stored energy at time period p is the stored energy in the previous time period $(p-1)$, plus the charged energy and minus the discharged energy, both at time period p. A more detailed mathematical representation of ESS could include energy losses due to charge, discharge, and self-discharge [44]. Here one important aspect arises: representing properly the ESS in an optimization model requires the preservation of chronological information. In other words, the sequence of time periods is highly important for the ESS representation. Furthermore, due to the different type of ESS technologies, this constraint could be representing ESS that store energy in minutes, hours, days, or months. For instance, Figure 2-2 shows a typical storage level for two different ESS technologies: BESS and a hydro reservoir. In this example, the BESS has completed almost seven full cycles[5] in one week, while the hydro reservoir only has completed almost one full cycle in one year (due to the seasonal behavior of this type of ESS).

[5] Full Cycle: The complete discharging and charging of a storage system.

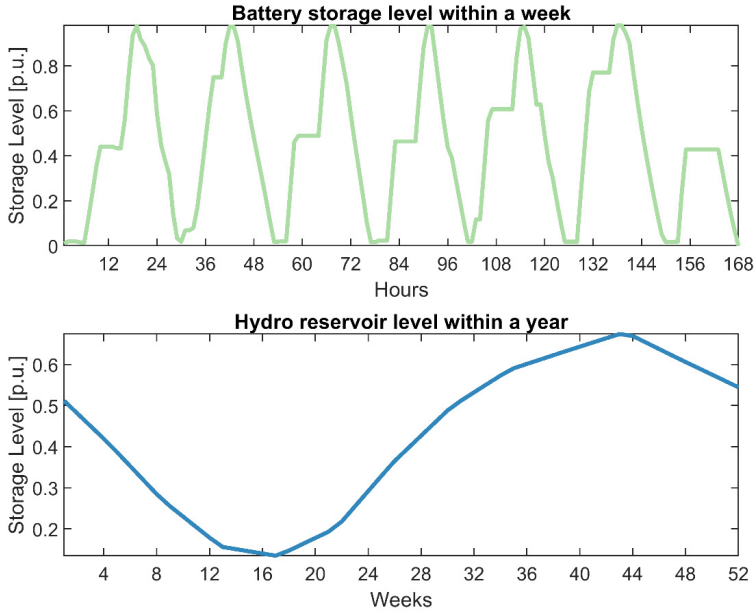


Figure 2-2. Example of energy storage level over time

On the one hand, the mathematical representation for the battery needs at least an hourly time period in the storage level equation. On the other hand, the hydro reservoir only needs a weekly or even monthly time period. This means that we need a mathematical representation of this technical constraint in which both type of storage LTESS and STESS can be properly considered at the same time.

A naïve approach could be solving a detailed model with the lower time resolution (e.g., hourly), however, this is not an easy task because MLTOP models solved using chronological hourly information can become computationally intensive for long-time horizon problems [45]. Instead, the practical approach is to use the Load Duration Curve (LDC) method [46], which groups together hours into load levels, a.k.a. load blocks, that have a predefined power and duration. Figure 2-3 shows an example with an hourly demand in a year and the correspondent LDC with five load levels. Instead of using the whole hours in the year to solve the model, the LDC method uses five load levels such that their durations sum 8760 h. Therefore, it is faster to solve the model using the LDC method than solving it with the hourly approach.

Despite the fact that the LDC method is more computationally efficient than the detailed hourly approach, chronology among individual hours is lost, and hence, chronological constraints cannot be represented. This information is necessary to properly model some technical constraints, such as the short-term operation of a storage facility [47]. The following section analyzes the main approaches in the literature to overcome this situation, including their major advantages and disadvantages.

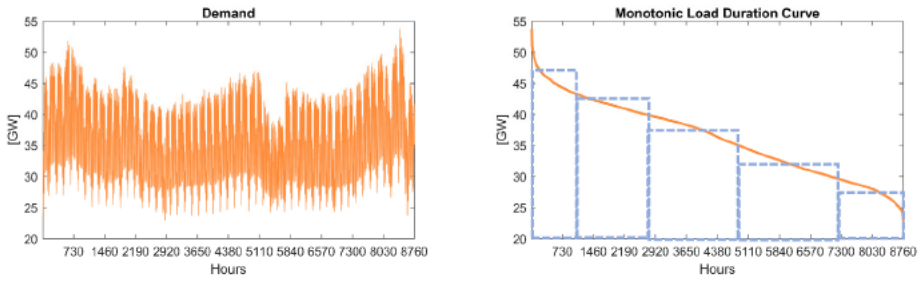


Figure 2-3. Example of Load Duration Curve

2.3 Common Approaches to Deal with Chronological Constraints

There are two common ways to reduce temporal information while maintaining some chronology that can be found in the literature: *representative periods* (RP) and *system states* (SS). The system states are also referred to as load levels, load duration curves, or time slices in more simplified versions. Both methods are based on clustering techniques in order to obtain the representative periods or the system states. The input data in the clustering process is usually hourly demand and renewable energy (e.g., wind and solar) time series. The main difference in the clustering process is that representative periods are constrained to a predefined length of the representative period, while the system states are not. To illustrate this point, Figure 2-4 and Figure 2-5 show an example of the input data in both approaches. On the one hand, Figure 2-4 shows the input data for demand and wind production as a 8760x2 matrix for one year, and then after the clustering process the result is N system states each one with its demand and wind production. On the other hand, Figure 2-5 shows the input data for demand and wind production as a 365x48 matrix for the same information, assuming a representative period equivalent to a one day. After the clustering process, M representative days are obtained in a Mx48 matrix with the information of demand and wind production for each representative day.

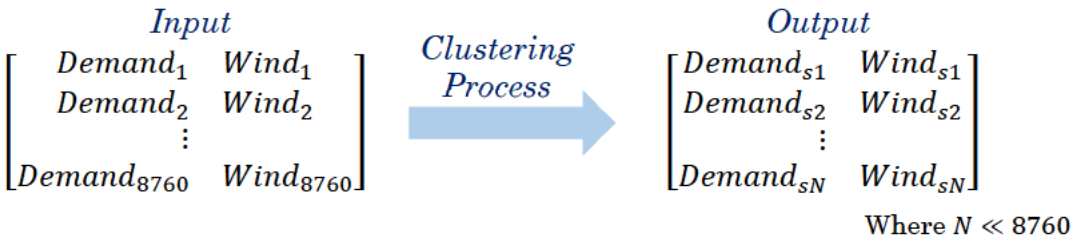


Figure 2-4. Input and output data format in the clustering process for SS

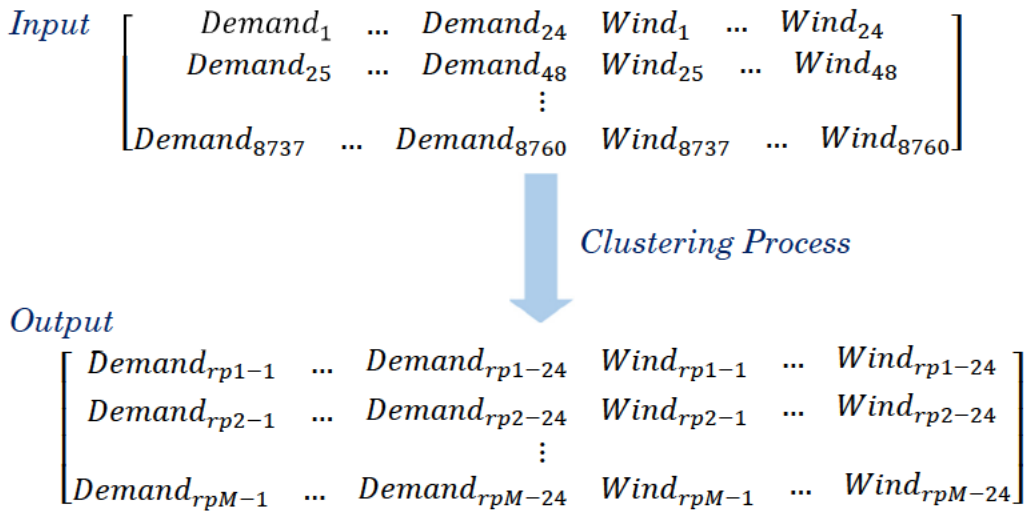


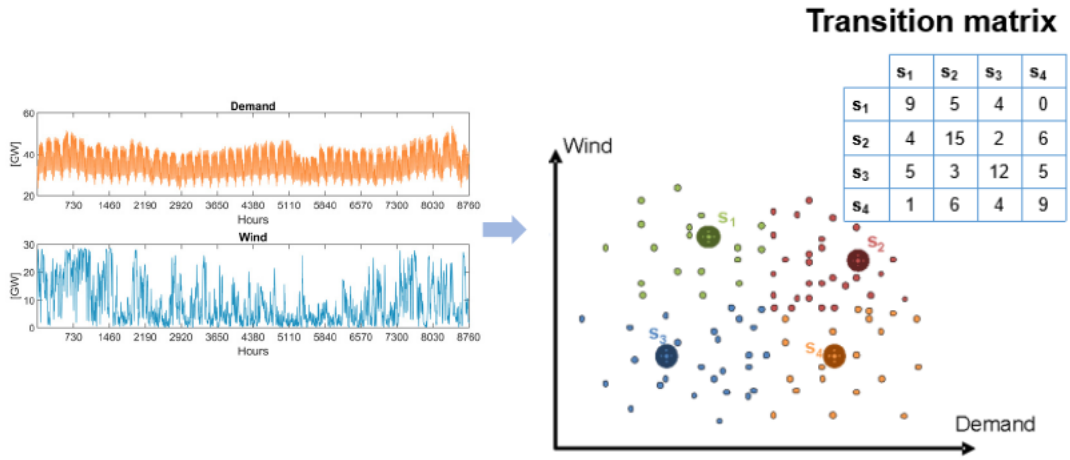
Figure 2-5. Input and output data format in the clustering process for RP

In this section, we show that the system states method is better than representative periods method from the point of view of LTESS, nevertheless, the representative method is better than the system states method from the point of view of STESS. In addition, we describe the main characteristics of both methods and review publications that present them for their application in medium- and long-term planning models.

2.3.1 System States

The *system states* approach, was introduced in [48]. It is designed to be an improvement on the entirely non-chronological LDC method. The SS method characterizes each time period (e.g., hours) in the time horizon by a set of features such as demand, wind, and solar power availability. For instance, if we consider hours as the time periods then hours with similar values of these features are considered to belong to the same *system state*, see Figure 2-6. Every hour in the time horizon is then assigned to one of the system states, $s = \{S_1, S_2, \dots, S_S\}$, and calculations are done for each system state in the same way they would be done for each hour of an hourly model. Each system state gets a weight or duration that depends on the number of real time periods in the time horizon that are represented by it. This is also called time slices in models such as ReEDS [49]. The innovation of SS method in [48] is the *transition matrix* ($N_{s,s'}$), where each element in the matrix represents the expected number of times that we have a transition from state s to state s', such as a transition matrix in a Markov chain. In other words, each number in $N_{s,s'}$ could translate to the number of times that we go from state s to either state s' during the whole-time horizon. For instance, Figure 2-7 shows four system states with the corresponding transitions among them. In this example, the element $N_{s2,s3} = 2$ indicates that during the whole-time horizon, two times, hours belonging to system state s_2 are followed by hours belonging to system state s_3. Note that this matrix does not necessarily have to be symmetric, since it depends on the chronological information of the hourly information (e.g., element $N_{s3,s2} = 3$). All this information is used in operational planning

models as chronological constraints, such as start-up constraints, which represents an improvement compared to the classic LDC method. These constraints are shown in Appendix A.3. However, the works based on SS framework have not considered the influence of transmission networks so far. The incorporation of the transmission network data into MLTOP models is important because congestion between different areas of the network could increase operational costs [11], [50].



	s_1	s_2	s_3	s_4
s_1	9	5	4	0
s_2	4	15	2	6
s_3	5	3	12	5
s_4	1	6	4	9

Figure 2-6. Example of SS definition by clustering method

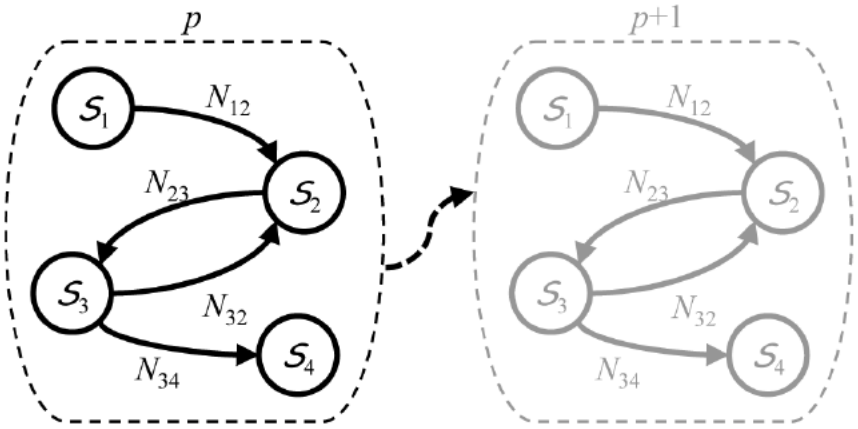


Figure 2-7. Transitions among system states [48]

The representation of ESS using the SS approach is not elementary. Since a system state represents a group of different hours in the year (continuing with time periods equal to hours), the definition of the storage level per system state would imply that all hours belonging to the system state have the same storage level. This is not necessarily true because if a system state includes hours in spring and some others from summer then the storage level of a hydro reservoir will hardly be the same for these hours that belong to the same system state. In [51], the authors propose an extension of the SS approach to tackle this problem. Instead of determining the storage level in each system state, the authors propose to calculate the change in stored energy between two system states. This proposal relies on the assumption that the change in the storage level between to system states is the same, no matter if each system state groups different hours. For instance, if the change in the stored level between S_1 and S_2 is positive, i.e., charging energy, then you can assume that each hour belonging to S_1 that is followed by an hour belonging

to S_2 implies a positive change in the storage level. Therefore, the total storage in any given hour can be calculated ex post by adding up all the changes in storage from the beginning of the time horizon to the hour of interest.

The total storage is kept within bounds during the modeling process by backtracking to calculate the total storage at certain chosen hours in the time horizon and constraining storage in those hours to be within bounds. This is achieved by introducing a *frequency matrix* in the formulation, which contains those hours where the storage must be within the bounds. In addition, an Incremental Bounding Algorithm (IBA) is used to determine the hours in the frequency matrix. For the sake of briefness, the frequency matrix and the IBA are not explained in this section, however, a more detailed explanation is given in Appendix A.3. The time to calculate and pre-select the hours in which the total storage must be in bounds is the main drawback of ESS representation using the SS method, especially for STESS, because several hours must be included to maintain its storage level within bounds. Moreover, as the number of hours in the frequency matrix increases it also increases the CPU time to solve the model. In Chapter 3, we show the implications of this situation in the approximation made by the SS method and propose a solution to improve this drawback on the representation of STESS. In contrast, LTESS only reaches its storage limits a few times in a week or month, making it easier to determine the hours in the frequency matrix. Therefore, LTESS are represented more accurately and efficiently in the SS method than the STESS.

2.3.2 Representative Periods

In the *representative periods*[6] method, a certain number of days, groups of days, or in some cases weeks that are representative of the variety of situations that can be found during the course of the time horizon (e.g., year) are chosen. All calculations (e.g., investment decisions and unit dispatch) are done for the selected days or weeks. Figure 2-8 shows three representative days selected from the hourly demand and wind time series in a year. Each *representative period* represents the periods in the year that are similar, so one can reconstruct the behavior of the system over the whole year by using the values calculated for the RPs in place of the periods they represent. The RP method preserves the internal chronology of the hours, yielding a more realistic representation of changing storage levels over the course of a day or week. However, the RP method does not preserve the chronology among the RPs. Therefore, any ESS with a full cycle longer than the representative period (e.g., weekly, monthly, or yearly rather than daily) will not be chronologically represented with the highest accuracy. This means that STESS is represented more accurately in the RP method than the LTESS. The RP method has also been used for some of the models that try to incorporate both long- and short-term dynamics, such as the RPM model in ref. [54]. However, these attempts use iterative approaches to integrate both dynamics, and therefore, they lose the possibility to co-optimize both types of energy storage systems.

[6] In this thesis, we use the name *representative periods* when general concepts and model formulation are explained. However, some authors call it *representative days* because the selected period is equal to a day [52], [53].

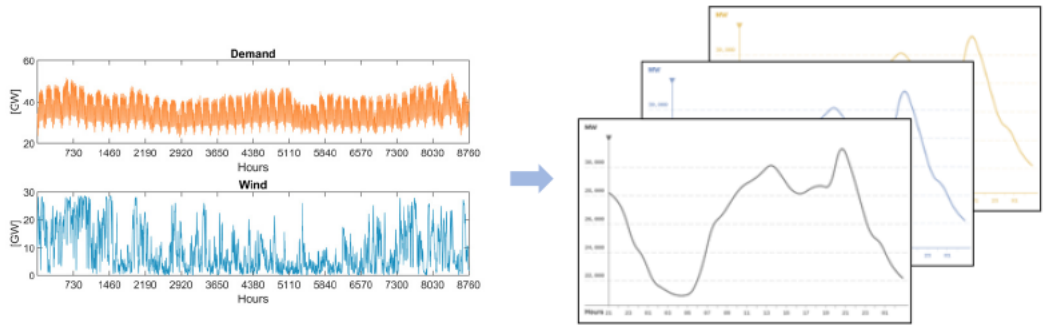


Figure 2-8. Example of RP definition by clustering method

There has been much research about the best way to choose these RPs. Some authors use a heuristic method, choosing one day for each season or one day in each season for the week and for the weekend. However, most of the methods rely on clustering techniques such as k-means [55], k-medoids [56], or hierarchical clustering [52], [57]. Others have proposed more tailored methods that involve optimizing the number of RPs that minimizes the difference between the load duration curve and the approximate one created by the selected RPs [53], [55]. There has also been research about the optimal length for RPs. For instance in [58], the authors suggested representative groups of days or representative weeks, whose advantage is that they increase the amount of chronology preserved, and whose disadvantage is that it increases the computational burden. The most versatile method for grouping RPs comes from Nahmmacher et al. [52], and relies on clustering techniques to group a number of hours with any number of normalized characteristics (solar energy, demand, wind energy, etc.) in network-constrained European case study using the LIMES-EU model. No matter how long the periods or how they are chosen, the drawback of this RP method is that it can only deal with relatively short-term storage cycles, those that charge and discharge in the course of a period (e.g., day), but not, for example, with seasonal storage (e.g., hydro reservoirs) with monthly or yearly cycles.

2.3.3 Challenges on Current Approaches

Table 2-1 shows the main advantages and disadvantages for each method analyzed. We can summarize this analysis by stating that SS method is better than RP method from the point of view of LTESS representation, nevertheless, the RP method is better than the SS method from the point of view of STESS. Therefore, the main challenge is how to combine them in order to obtain the best of both methods. This idea is explored in Chapter 3 to propose a new method that can handle both types of ESS and, therefore, co-optimizing both operational and investment decisions. In addition, the Chapter 3 analyzes the operation of a network-constrained power system using the SS method, considering that this is one of the identified drawbacks for this method.

Table 2-1. ESS representation in SS and RP methods

Method	Advantage	Disadvantage
System States (SS) [51]	The representation at the same time of short-term ESS and long-term ESS.	The definition of ESS bounds before solving the complete model, which hinders the accurate representation of short-term ESS.
Representative Periods (RP) [52]	The accurate representation of short-term ESS within the RP.	ESS that can store energy for a time period longer than the RP, i.e., long-term ESS, cannot be represented and normally its production is predefined within the RP.

2.4 Generic Formulation of the Unit Commitment Problem

The UC problem is widely used for a System Operator[7] (SO) to schedule the generation resources in order to achieve an economical energy production in the power systems subject to technical constraints of these resources [59]. Recent studies [60], [61] include the UC formulation in medium- and long-term planning to correctly represent the power system operation and its main constraints. For instance, authors in [62]–[64] have shown the importance of including short-term dynamics on strategic planning (i.e., investment decisions) in order to consider the increased need of operational flexibility due to VRES integration. In this thesis, the UC formulation is used as the base set of constraints to perform the analysis of operational and investment decisions. Here we show a compact matrix formulation based on [65]:

$$\min_{\boldsymbol{x},\boldsymbol{p},\boldsymbol{r},\boldsymbol{c},\boldsymbol{s}} \ \mathbf{a}^{\intercal}\boldsymbol{x} + \boldsymbol{b}^{\intercal}\boldsymbol{p} + \mathbf{d}^{\intercal}\boldsymbol{r}$$

$$s.t. \quad \mathbf{F}\boldsymbol{x} \leq \mathbf{f},\ \boldsymbol{x}\ is\ binary \tag{2-1}$$

$$\mathbf{H}\boldsymbol{p} + \mathbf{J}\boldsymbol{r} \ \leq \mathbf{h} \tag{2-2}$$

$$\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{p} \leq \mathbf{g} \tag{2-3}$$

$$\boldsymbol{r} \leq \mathbf{R} \tag{2-4}$$

$$\mathbf{D}\boldsymbol{p} + \mathbf{E}\boldsymbol{c} + \mathbf{K}\boldsymbol{s} \leq \mathbf{e} \tag{2-5}$$

where $\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{c}$, and $\boldsymbol{s}$ are decision variables. The binary variable $\boldsymbol{x}$ is a vector with on/off status and startup/shutdown decisions for each generation unit and time interval (e.g., hours) over the whole planning horizon. The continuous variable $\boldsymbol{p}$ is a vector for each unit production decision and time interval, including ESS (i.e., discharge decisions). The continuous variable $\boldsymbol{r}$ is a vector of renewable production decision (e.g., wind, solar productions) for each time interval. The continuous variable $\boldsymbol{c}$ is a vector of charging decisions for each

[7] Depending on the regulation of the country, SO could be an independent organization (Independent System Operator – ISO) or part of the transmission company (Transmission System Operator – TSO).

ESS and time interval. The continuous variable $\boldsymbol{s}$ is a vector of stored energy for each ESS and time interval.

The objective function is to minimize the sum of the commitment cost $\boldsymbol{a}^{\intercal}\boldsymbol{x}$ (i.e., non-load, start-up and shut-down costs), production cost $\mathbf{b}^{\intercal}\boldsymbol{p}$, and renewable production cost $\mathbf{d}^{\intercal}\boldsymbol{r}$ over the planning horizon. ESS production cost and renewable production cost are usually considered to be zero. However, their cost parameters are explicitly included to consider the possibility where these costs are different than zero. In both cases, these parameters could represent operation and maintenance cost.

Equation (2-1) contains only commitment-related variables, e.g., minimum up and down times, startup and shutdown constraints, variable startup costs. Equation (2-2) includes production-related constraints, e.g., energy balance, reserve requirements, transmission limits, ramping constraints. Equation (2-3) combines the commitment and dispatch decisions for thermal generation units, e.g., minimum and maximum generation capacity constraints. Equation (2-4) states that renewable production cannot exceed its predicted values **R**. Equation (2-5) defines the ESS constraints, e.g., maximum charge/discharge capacity, maximum storage capacity, and storage balance constraints.

Appendices A to D detail all the models used in this thesis. Operational decisions in these models are based on the described UC generic formulation in this section.

2.4.1 Flexibility Requirements in Power Systems and the UC

Among the strategic planning models, Generation Expansion Planning (GEP) is one of the classic long-term problems in power systems that aims at determining the optimal generation technology mix [66]. Environmental policies, such as renewable targets [67] or CO2 emission reduction [68] influence in GEP decisions, leading to the integration of vast amounts of Variable Renewable Energy Sources (VRES), i.e., wind and solar, in GEP. Nevertheless, VRES integration has consequences in GEP modeling. For instance, previous studies [62]–[64] have shown the importance of including short-term dynamics on GEP decisions in order to consider the increased need for operational flexibility due to VRES integration. Therefore, correctly modeling flexibility in GEP models is crucial to reach the right conclusions in the energy transition process.

In order to consider operational flexibility in GEP, Unit Commitment (UC) modeling is needed to determine system operation [64], [60]. For example, it is known that units are being cycled more frequently due to higher VRES flexibility requirements [69]. Studies have shown that ignoring startup and shutdown processes highly overestimates the flexibility and underestimate the costs of the system [31]. Another example is ramping constraints. If we focus on flexibility and want to know a good (optimal) future generation-mix and interconnection capacities for a given scenario, the GEP problem must include at least detailed ramping constraints. Moreover, operating reserve decisions have also become more relevant in GEP with the integration of VRES because they may ensure that generation technologies have an extra income to recover their investment costs through these types of ancillary services.

Despite the recent developments to consider flexibility requirements in GEP, classic GEP models are proposed using energy-based UC models. The following section discusses the implications of energy-based UC models to consider the flexibility requirements in power systems. Moreover, it compares the energy-based UC model with its alternative, the power-based UC model.

2.4.2 Energy- vs Power-based UC for Flexibility Requirements

Power[8] and energy[9] are concepts that are tied to each other, and yet they may have a different impact on the optimal scheduling of generation units. For instance, Figure 2-9 shows the power and energy values of demand in two consecutive hours. The energy value of demand in both hours is the same (15 MWh), while the power value of demand at the end of each hour is 20 MW and 10 MW respectively. Therefore, if we consider energy value of demand to schedule the generation units, the load ramp is zero. However, this is not true, since the power values of demand show that the upwards and downwards ramps are 10 MW/h. Thus, scheduling decisions in the UC problem may change depending on whether they were taken using one concept or the other. Nevertheless, the power-based schedule will be better prepared than the energy-based schedule because the latter cannot capture the variation of the demand during the hour through the load ramp. This is something that system operators deal with in real-time operation by using secondary reserve or rescheduling generation units within the hour.

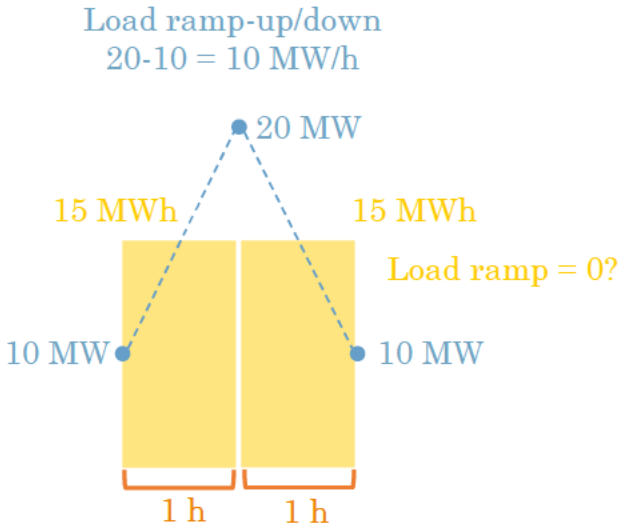


Figure 2-9. Differences between energy and power

Recent studies [31], [32], [70] have shown that energy-based UC models cannot capture variability on demand, and cannot deliver the flexibility that is expected (i.e., overestimate the flexibility of the system). This is mainly because average energy levels (e.g., average level in one hour) do not provide detailed information about the instantaneous output of a generator, and con-

[8] Power (MW): Instantaneous value of electric demand, generation production, or transmission flow

[9] Energy (MWh): the accumulative value of power in a period of time (e.g., 1h). It is proportional to the average value of power in that period of time.

straints such as ramping-limits and demand-balance are dependent on instantaneous outputs rather than average levels. This means that more flexibility than planned by energy-based models can be used and thus is used in real-time operation (through operating reserves and allowing deviations on schedules) in order to deal with all the problems introduced by these traditional energy-based models.

Power-based models have been proposed [71] to overcome the energy-based model flaws. This is achieved by better representing the exploitation of system flexibility [31] and allowing the correct modeling of operating reserves and ramping constraints [32], [70] in order to deliver the expected flexibility from the generation resources. This is possible because a power-based model has a clear distinction between power and energy in its formulation. Demand and generation are modeled as hourly piecewise-linear functions representing their instantaneous power trajectories. The schedule of a generating unit output is no longer an energy stepwise function, but a smoother piece-wise power function. Figure 2-10 shows an example of demand in both models.

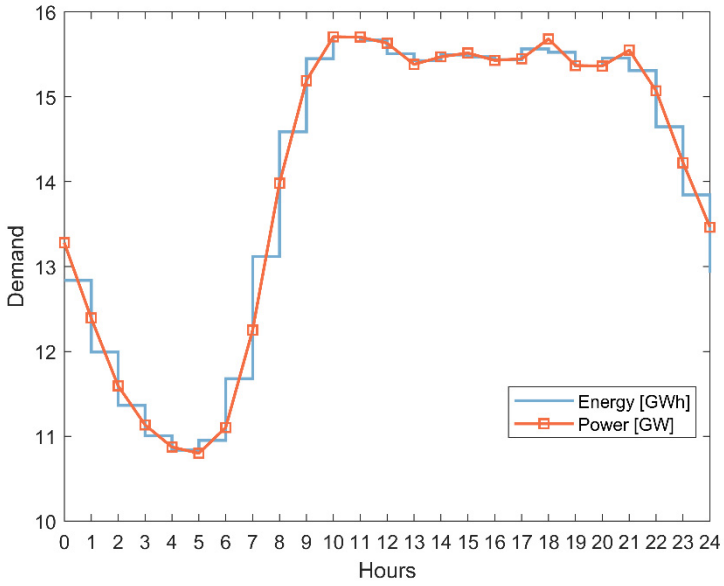


Figure 2-10. Power and energy demand

Another important aspect to determine the flexibility requirements in power systems is time resolution. In order to model correctly the real operation of power systems a high resolution is needed (e.g., minutes). Current GEP models are based on hourly resolution where the underlying assumption is that it is enough to capture the variability and flexibility requirements of power systems. However, reducing the time resolution (e.g., from hours to minutes) in long-term investment models would make them almost impossible to solve. Instead, it has already been shown in [31] that real-time simulations (e.g., 5-min time step) help to determine the performance of different schedules (operational decisions) to meet the real-time flexibility requirements in the power system. This type of real-time validation is not common to be carried out because it is considered unnecessary. Nevertheless, to validate correctly flexibility capabilities and requirements of the system, this real-time evaluation is paramount [31].

Chapter 6 shows a detailed comparison among the different constraints using either energy- or power-based formulation.

2.4.3 Challenges on Power-based UC Model to Represent ESS

Despite the recent developments and research in power-based UC formulations, ESS have not been integrated in this formulation so far. Moreover, ESS are one of the most promising options to provide flexibility in power systems [41], and the representation of flexibility requirements is one of the main advantages in power-based UC formulations. However, the main challenge here is to exploit correctly the ESS capabilities to provide flexibility through a careful modeling of ESS using power-based formulations (ramping and reserve constraints). Chapter 6 shows a proposed model for strategic planning in which ESS investment and operational decisions are considered using a power-based formulation.

2.5 Energy Storage Regulation Background

So far, we have analyzed technical aspects of ESS modeling in tactical and strategic planning models, which is the focus of this thesis. Nevertheless, another important aspect is the regulatory framework of ESS technologies. Therefore, here we take a quick look on policies related to ESS in order to add a complementary point of view on ESS topic in the context of this thesis.

The applications of ESS will depend on the power system characteristics, their own characteristics, and on the type of VRES installed. As we have seen in this chapter, ESS can play different services in the power system. Therefore, one question arises: are the ESS generators, loads, or transmission/distribution assets? The answer to this question leads to a regulatory debate, i.e., whether ESS should be considered as network assets, as generation assets, or as a new separate asset category [72]. On the one hand, if ESS are considered as generation assets, then unbundling conditions are needed to prevent network businesses (i.e., natural monopolies) from owning and operating ESS in liberalized activities. On the other hand, if ESS are classified as network assets, then they must provide network services only, i.e., avoid participating in liberalized activities. Therefore, given the diverse roles that ESS can play, some authors have even suggested that ESS should be considered as a new type of asset to solve this dilemma [72], [73]. Moreover, Conejo and Sioshansi [74] analyze the major challenges in designing electricity markets to embrace new technologies that provide opportunities for a more active participation by consumers, including those related to ESS and distributed energy sources. These challenges show the need for new design principles for electricity markets in order to regulate the ESS technical operation while guaranteeing a proper remuneration of their services. The complete analysis and discussion in this section were published in Article IV [10].

2.5.1 ESS Policies in the US

US policies can be divided into state and federal jurisdictions. At the state level in recent years, several states have introduced policies aiming to support the integration of ESS in electricity markets. Some states have included ESS in their energy capacity planning, creating specific programs and even cofunding some projects [40]. However, these policies at the state level show a

lack of a common approach in the US for ESS deployment. Each state proposes rules depending on its own priorities to incentivize utility-scale or distributed ESS. This situation explains why ESS have thrived in some states and not in others [75], [76]. At the Federal Energy Regulatory Commission (FERC) jurisdictional level, the PJM system is one successful case in the US for ESS integration [77]. In PJM's wholesale markets, ESS can participate in energy, capacity, and ancillary service markets. Pumped-hydro storage participates in all these markets; however, battery and flywheel storage technologies participate only in regulation markets (i.e., ancillary services) providing fast regulation service. The main reason for this situation is that battery and flywheel owners have enough economic signals from the reserve markets without the risk of penalization in the capacity market. Nevertheless, this situation could change due to recent federal rules.

On 15 February 2018, the FERC published the Order (Order 841) [78] to integrate ESS more effectively into wholesale markets in order to enhance competition with proper economic signals. The Order 841 derives from concerns regarding the barriers that ESS may face, which would hinder their participation in organized wholesale electric markets. Three key challenges can be drawn from Order 841: the participation models for ESS in the security-constrained unit commitment, economic evaluation, and regulatory treatment (i.e., ownership).

First, Order 841 establishes that ISOs must represent the physical and operating characteristics of ESS through bidding parameters or other means. FERC includes the following parameters in this bidding format: charging/discharging limits, rates, times, and run time, as well as the state of charge (SoC). These bidding parameters will allow the ISOs to optimize ESS dispatch more efficiently. Moreover, ESS agents should have the option of self-managing the SoC using this bidding format. This option offers to ESS agents the possibility of providing multiple services in the power system. However, Order 841 supports the idea that the ESS is more efficiently dispatched when it is in the hands of the system operator.

Second, the economic evaluation of ESS needs a wider perspective. Therefore, the Order establishes that ESS is eligible to provide all services (e.g., capacity, energy, and ancillary services) that the resource is technically capable of providing. As a result, ESS could find different revenue streams to leverage their investment. Nevertheless, ESS could be still expensive to provide some services in the power system (e.g., as an alternative to peaking plants with fast capabilities). In addition, ESS enable the integration of a high VRES proportion, and they should be properly compensated for these benefits in order to guarantee their cost recovery. Other mechanisms such as forward capacity markets should be adapted to enable the participation of ESS, e.g., allowing them to be aggregated with renewables sources, demand response, or energy efficiency.

Third, FERC does not explicitly mention rules regarding ESS ownership. Recently, FERC issued a policy statement [79] in which the scenario of ESS as a transmission asset is analyzed. This statement mentions that there is no regulatory impediment for ESS to provide transmission and generation services at the same time. However, several concerns arise in this scenario. For

instance, RTO/ISO independence and double recovery of costs are among the main concerns. In order to tackle these concerns, the California Public Utilities Commission (CPUC) issued a decision on multiple-use application issues [80], which provides direction to the utilities on how to promote the ability of ESS to realize their full economic value when they can provide multiple benefits and services to the electricity system. This decision defines eleven rules to determine the evaluation of these multiple-use ESS applications, as well as definitions of service domains, reliability services, and non-reliability services. Nevertheless, this decision still leaves some open issues such as tariffs, aggregation with distributed energy resources, appropriate metering, measurement, and accounting methodologies. Therefore, the discussion on ESS ownership versus the provision of multiple services is still an open topic, and its resolution will condition future ESS deployments in the US.

2.5.2 ESS Policies in the EU

At the European Union level, the electricity industry is regulated by The Electricity Directive - Directive 2009/72/EC and The Renewable Energy Directive - Directive 2009/28/EC. These Directives aim for the completion of the Target Model for the Single Energy Market for Europe. There are many references to electricity storage in the existing regulation. However, further details are required. For instance, The Electricity Directive includes a list of definitions regarding power generation, transmission, distribution, and supply terms. Nevertheless, the concept of ESS is not mentioned in the document. The Directive fails to include ESS as a separate component in the electricity sector structure. As result, ESS is generally treated as a generation asset in Member States [81].

This situation is changing in the Commission's *"Clean Energy for All Europeans"* proposals [82] and, particularly, with an improved regulatory framework proposed under the Market Design Initiative (MDI). For instance, the following definition of energy storage is included: *"Energy storage in the electricity system would be defined as the act of deferring an amount of the energy that was generated to the moment of use, either as final energy or converted into another energy carrier."* However, ESS is not established as a separate component of the power system with its own characteristics, and this could restrict the potential of ESS [83]. The proposal also removes discriminatory network tariffs (e.g., double grid fees) that unnecessarily disadvantage ESS.

The development and operation of storage facilities is promoted in the new MDI as a commercial activity to be performed by market participants rather than regulated entities. TSOs and DSOs should not own, manage, or operate ESS facilities. In exceptional cases, the system operators could be allowed to invest in an ESS facility under regulatory approval and supervision only if other market parties are not interested in providing a specific ESS service. According to the EU [84], in these cases, the regulatory authorities should regularly reassess the potential interest of market parties to be involved in such activity.

In February 2017, alongside the Second State of the Energy Union report, the European Commission published a Staff Working Document entitled: *"Energy storage – the role of electricity"* [84]. This document outlines the role of energy

storage in relation to electricity, presents the advantages of different technologies and innovative solutions in different contexts, and discusses possible policy approaches. In summary, the development and financing of ESS should depend on the following principles:

- ESS should be developed to the extent that the overall costs of the new power system are lower with storage than without storage.
- In relation to the electricity grid, ESS should be rewarded for the services provided with alternative suppliers for those services, either demand response or flexible generation.
- The supporting role of ESS in integrating VRES should be rewarded for its contribution to improved energy security and electricity sector decarbonization. In addition, the avoided costs of VRES curtailment and the carbon reductions could also support the business case for large-scale ESS.
- If either a consumer or a generator wants to integrate an ESS at its current facilities, then this should not lead to less favorable treatment (e.g., discriminatory grid access, or paying at the same time grid fees as both consumer and producer) either in terms of obligations or in terms of eventual support that it receives in the power system.

The EU is addressing these principles for ESS by promoting innovation in key technologies and developing suitable market rules. Technological innovation in storage falls under the Horizon 2020 programme [85] and the Strategic Energy Technology Plan [86]. Moreover, large storage projects above 225 MW are included in the selection process for the EU's projects of common interest (PCI).

2.5.3 Challenges on ESS Regulation

From the regulatory point of view and according to the reviewed legislation in earlier sections, it is possible to summarize the key challenges on ESS regulation into two: (i) the regulation of the ownership of storage to avoid an outcome with insufficient unbundling, which may hamper market operations, among other considerations; and (ii) the need to rethink market design across timeframes (i.e., capacity, day-ahead, intra-day, and real-time markets).

Regarding ESS ownership, unbundling principles forbid its ownership by regulated entities. Nevertheless, in the particular case of ESS, this leads to an inefficient realization of the full ESS potential. By contrast, allowing ESS ownership by regulated entities (i.e., TSOs or DSOs) may enable ESS to provide network services; however, it may create a conflict of interest or market inefficiencies due to the monopoly nature of these entities. Therefore, the crucial regulatory challenge is to guarantee that ESS can provide market and network services as well as market efficiency. This efficiency of market mechanisms could be made possible by eliminating cross subsidies between regulated and market parties and avoiding conflicts of interest. As a possible solution, some authors [83], [87] have proposed allowing grid operators to procure system flexibility services from third-party ESS operators in the market. The creation of a proper market for ESS services could mitigate concerns about ESS ownership. In addition, more competition could be introduced to this market if small players are allowed to participate, individually or through aggregation. If properly implemented, then this reform could also

address issues related to the provision of cost- and market-based services. A third-party ESS provider has advantages because TSOs or DSOs could use competitive offers to obtain network services and, through their bids, incorporate potential revenues from market-based services that are unrelated to network services. Therefore, the ESS owner could deliver network services and participate in the markets (e.g., DAM, intraday, or balancing), every time the TSO or DSO has not contracted the ESS services, and incomes (or penalizations) from the operation of the ESS in the wholesale market would belong to the third party and liberalized owner of the ESS. In contrast to a third-party ESS provider, Sioshansi [88] has proposed a solution where storage-capacity rights are auctioned to third parties that use their rights for cost- or market-based services. As in the third-party ESS provider proposal, the benefits that the storage asset provides are separated from the regulatory treatment of those benefits (e.g., either competitively priced or unpriced), guaranteeing that ESS assets can recover their cost. A special characteristic of storage-capacity rights is that they are agnostic to who operates the storage capacity auction. Therefore, even an ISO may be able to operate the auction without threatening its market independence. Authors in [41] show the increase in the commercial value derived from ESS provision of network and market services. This is possible only in a regulatory framework that balances synergies and conflicts among the provision of different types of services while maximizing ESS revenues.

Regarding market design, it was shown that ESS are eligible to participate in DAMs of both the EU and the US. However, the inter-temporal constraints of ESS provide challenges to guarantee that the ESS is scheduled within their operational parameters in the most efficient way. The bidding options for ESS in DAMs should give market signals for flexibility in the power system. On the one hand, the EU approach does not guarantee the most efficient operation of ESS because the linked block order limits the charging/discharging hours to some predefined values that cannot be optimized in the DAM in order to increase total system welfare. On the other hand, this is different from the FERC approach, which suggests that ISOs could more efficiently optimize their dispatch. However, the FERC approach is suitable only if there are no market failures (e.g., lack of competition) and market rules are fulfilled; otherwise, it could hamper ESS development if there are market failures such that the efficient dispatch, performed by the ISO, does not allow ESS to obtain sufficient revenues on their investments. The lack of market signals makes it challenging for an ESS investor to make a business case for deployment. In fact, the authors in [89] have shown in the EU context that the revenues of ESS performing arbitrage in the DAM horizon are far from ensuring profitability in different markets. The authors of [73] state: “when the electricity market is well conceived, it remunerates correctly the services valuable to the electric system (e.g., capacity, energy, congestion management, real time balancing and frequency regulation) and it internalizes externalities such as congestion in nodal or zonal pricing of electricity.” In conclusion, current DAM rules in the US and the EU should be adapted to enable ESS participation in both network services and liberalized activities in order to obtain an optimal integration of these resources in power systems.

Outside the US and the EU, an example from Chile may provide a pathway for ESS in terms of regulation and new opportunities. The Chilean case is interesting because, for the first time in that country, the law 20.936(2016) [90] explicitly defined ESS as a power system asset, which is different from the existing definitions of generation and transmission assets. This opens the door to a wider possibility for integrating ESS properly with different kinds of services. Although the current definition allows ESS to participate only in the energy market, new regulation is under development to define the participation of this new asset in ancillary and network services. This could be a litmus test for future regulatory developments integrating ESS into power systems combining both liberalized market and network services. Both the US and EU market structures could benefit from this approach, which addresses the two key issues for ESS mentioned at the beginning of this discussion: ownership and market design. As a new asset, ESS should reduce the risk of insufficient unbundling for its owners in the market because they should be third party apart from generation and transmission activities. In addition, new market rules can be developed for this new asset, especially for situations when it provides part of its capacity for a network service (e.g., congestion management) and the remaining part in the liberalized markets (e.g., DAM, intraday, balancing, or capacity markets). Apart from the new asset approach, there has been discussions that focus on services that can be provided rather than the asset definition [91]. This approach also aims to unlock the ownership dilemma and market design issue by stacking multiple services that can lever ESS investment. In addition, focusing on services might provide other technological solutions such as aggregation of distributed generation, demand response, and distributed ESS. No matter the approach (i.e., either a new asset or new services), in both the US and EU frameworks, the major challenge is making new rules efficient enough that ESS owners have the right incentives to participate in both kinds of services while they recover the ESS investment without support mechanisms or subsidies.

Finally, it is important to mention that technology costs are currently the greatest barrier preventing further development of ESS. Despite recent cost reductions, ESS are still far from being treated as an economically competitive technology, although there are exceptions for particular uses, such as frequency regulation in PJM [77], integration of renewables in Australia [38], grid-balancing services in the UK [92], and transmission congestion management in Italy [93]. As a consequence, some R&D is still needed, and, as it is usually the case with immature technologies.

3 ESS OPERATION MODELING

Contents

This chapter studies optimization models for ESS operation in medium- and long-term planning. The results are divided in two main sections. The first section analyzes the impact of the transmission network in the clustering methods used to reduce the temporal information while conserving chronological information for the ESS. We analyze two main options to define the input data in the clustering process for this reduction: net demand per node, and demand and renewable production per node. These options are tested in a network-constrained case study in order to draw conclusions and recommendations for the representation of ESS operation. The second section compares the system states and representative periods methods when short-term energy storage and seasonal storage are included in the optimization models. In addition, enhanced versions of each method are proposed in order to overcome the main drawbacks in their former methods. Our results show that, on the one hand, system state models provide an excellent representation of long-term storage, however, their representation of short-term storage is frequently unrealistic. On the other hand, we proposed a new model called RP-TM&CI, which is one of the main contributions of this thesis. This model succeeds in approximating both short- and long-term storage by combining ideas from system states and representative periods.

The analysis and results in this section were published in Article I [94] and in Article II [95].

3.1 Impact of Transmission Network in Clustering Methods

This section analyzes the impact of including a transmission network on the clustering methods used to reduce the temporal information while conserving chronological information for the ESS in comparison with the single node case. In order to do so, we define two alternative sets of representative variables for clustering hours when the transmission network is considered. The analyses are made on the system states modeling approach.

3.1.1 Optimization Models

As described in Section 2.4, a UC formulation is used to represent the operation in power system including ESS. In this section, an hourly UC model is used as a benchmark, while a system states (SS) model including transmission constraints and ESS is analyzed with different clustering processes. Appendices A.2 y A.3 show both models in detail. Figure 3-1 shows an overview of the analysis. On the one hand, hourly input data is used to run the hourly model including the transmission network topology in order to obtain the hourly benchmark results. On the other hand, the hourly time series are clustered in order to obtain the system states. These system states are used to solve the system states model. The results are obtained for each system state. Then, it is possible to reproduce the hourly results considering that as an output of the clustering process the relationship among hours and system states is known. Therefore, it is possible to determine the quality of the approximation by analyzing both hourly results. This quality is measured by calculating the error of the hourly results in the system states model in comparison with the full hourly model, allowing to determine the best clustering process for each case study.

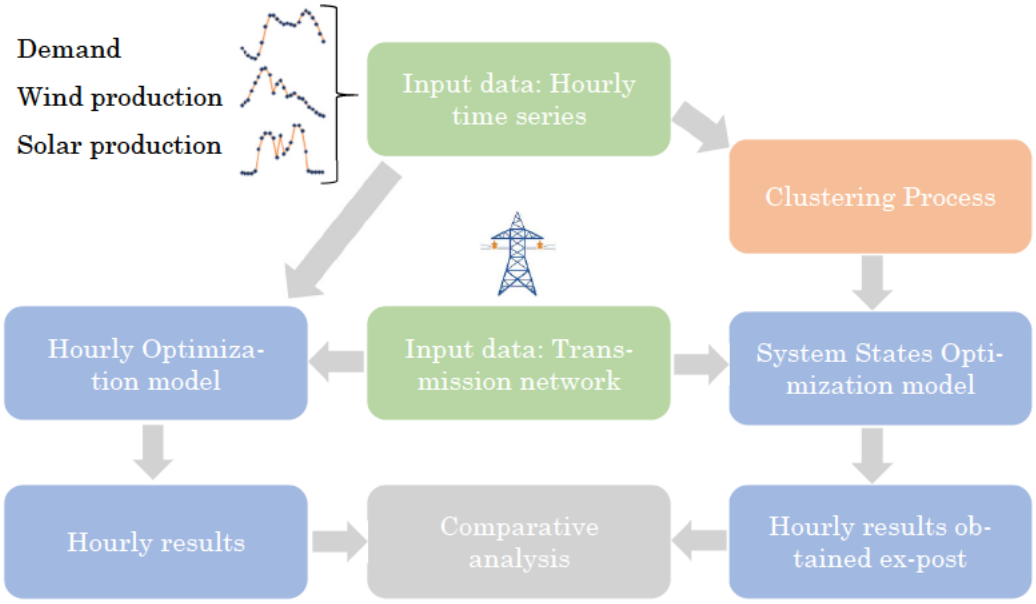


Figure 3-1. Analysis of transmission network in the SS model

3.1.2 Clustering Process in Transmission Constrained Networks

In order to determine the system states for different features of the system (e.g., demand time series, wind or solar production, or network topology), pattern recognition techniques such as Principal Components Analysis (PCA), Decision Tree (DT), or Cluster Analysis (CA) can be applied. PCA and DT approaches can be used first to obtain a dimension reduction in the features of the system. Once the main features are identified, a clustering algorithm (e.g., k-means++ or k-medoids [96]) can be applied to reduce the amount of elements at each feature. The goal of the clustering algorithm is to divide a set of features into a predefined number of subsets or *system states* so that the subsets minimize the sum of distances between a feature and the center of the measurement's cluster or *system state*. After the clustering-process has finished and the states have been defined, we can determine the *transition matrix* by calculating the number of changes from one state to another state considering the chronological sequence.

For a single node case, the information per hour of each feature is needed, i.e., demand, renewable production, or net demand. Figure 3-2 (left) shows an example using the feature net demand (ND_h) at each hour $h = 1, \dots, H$. Each system state has its centroid $\overline{ND}_s$. In this thesis, we extend this definition from a single node system to multiple nodes in a network. The main difference is the requirement to include spatial dimensions (i.e., bus or node $n = 1, \dots, \mathrm{N}$) in the clustering process. Consequently, a vector with net demand per hour at each node ($ND_{h,n}$) is required as input data. In this situation, each system state is a vector $1 \times \mathrm{N}$ (i.e., one centroid for each node). The clustering process minimizes the Squared Euclidian Distance (SED) between the information at each hour and the system states, and each hour is assigned to only one system state. Figure 3-2 (right) shows an example for clustering with a net demand feature. However, information such as demand and renewable production (instead of net demand) per node could be also used in the clustering process. The dimension of the input data vector for each hour is $1 \times 2\mathrm{N}$, therefore, each system state will be a vector $1 \times 2\mathrm{N}$. In Section 3.1.4, we analyze two ways of selecting the features for clustering and we draw some conclusions about them. The first one uses the dimensions of demand and wind per node (SS-DW) whereas the second one only uses net demand (demand minus wind) per node (SS-ND).

Finally, it is important to highlight that, the *system-state framework* attempts to reduce the information on time dimension and while it is not trying to reduce the spatial dimension (i.e., reduction of network nodes), such as other authors have studied [97], [98]. Further research could be addressed in order to combine both types of reductions techniques, i.e., temporal and spatial.

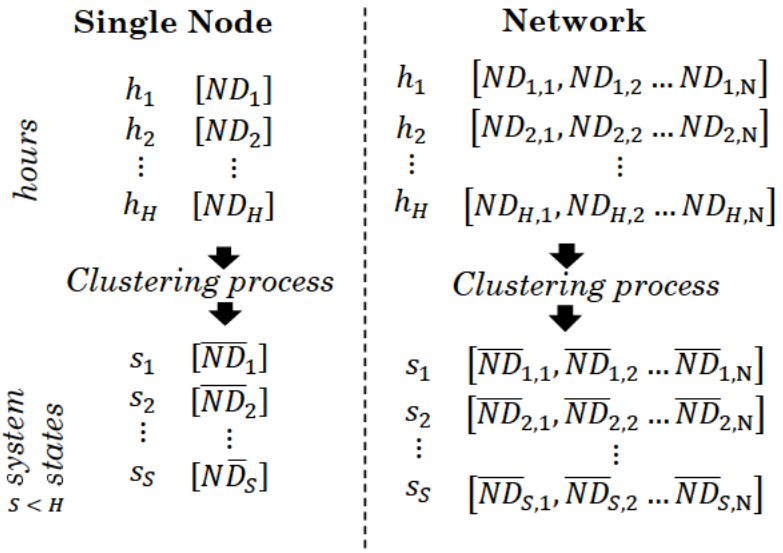


Figure 3-2. Clustering process for SS definition

3.1.3 Case Study Data

The network data used in this section is based on the 14-bus IEEE benchmark system, see Figure 3-3. The maximum transmission capacity is assumed to be 1500 MW for all lines and transformers in order to adapt it to the new demand and wind data. Based on [51], 11 different thermal generators were considered: 1 nuclear (NU) unit, 4 different types of coal (CO) units, 3 identical combined-cycle gas turbine (CC) units, 2 identical fuel-oil (FO) units, and 1 gas turbine (GT) unit. The models also consider two different storage technologies: a large long-term hydro reservoir at bus 1 and a fast-ramping BESS at bus 2. Efficiency data for the storage units are assumed taking into account the values shown in [99]. It is also assumed that the BESS is initially empty, and that the minimum final storage level is zero. The initial storage level of the hydro reservoir is 3000 GWh and the final level is required to be above 3000 GWh. Inflows are not considered for neither the hydro reservoir nor the BESS. These assumptions attempt to ensure that only energy arbitrage is driving changes in storage levels between the hydro reservoir and the BESS, as opposed to purely charging and discharging just to meet the initial and final level requirements.

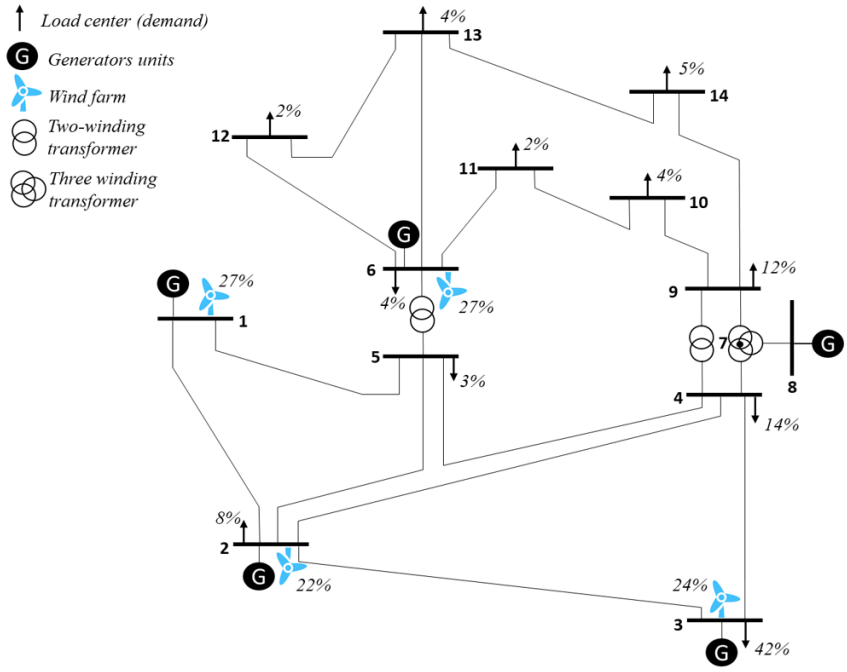


Figure 3-3. Modified 14-bus IEEE test system

3.1.4 System States Clustering in the Case Study

For the system states model, 100 different states were defined, two of which correspond to the first and last hour of the time horizon. The remaining 98 clusters were obtained via the clustering algorithm k-means++[10]. Section 3.1.8 shows a sensitivity analysis changing the number of states to validate the initial number of selected clusters. As mentioned in Section 3.1.2, two ways of selecting the features for clustering are analyzed: 1) demand and wind per node (SS-DW), and 2) net demand (demand minus wind) per node (SS-ND). These two ways of clustering are compared to the hourly model in the remainder of this section.

3.1.5 General Case Study Results

In order to compare the results of both models, two error metrics are defined. First, the *normalized root mean squared error* (NRMSE) defined in equation (3-1). The NRMSE measures how well of a fit the model is at each period, i.e., hour (power error). Second, the *accumulated error* (AE) defined in equation (3-2). The AE measures how well a model fits over all time horizons (energy error).

$$NRMSE = \frac{\sqrt{\frac{\sum_p \left(x_p - x_p^{ss}\right)^2}{N}}}{X^{UB} - X^{LB}} \times 100\% \quad (3\text{-}1)$$

$$AE = \frac{\left|\sum_p x_p - \sum_{ss} T_s \cdot x_{ss}\right|}{\sum_p x_p} \times 100\% \quad (3\text{-}2)$$

Where:

N: Total number of periods (e.g., hours)

[10] 10000 replications have been performed to select the 98 clusters. At the end of the clustering process, the 98 clusters that have the minimum distance are selected (i.e., squared Euclidean distance) from the 10000 replications.

x_p: result from the hourly model per period p (e.g., production)

x_{ss}: result from the system states model at each state ss

x_p^{ss}: result from the system states model associated to period p

X^{UB}: Upper bound of variable associated to result x_p

X^{LB}: Lower bound of variable associated to result x_p

The optimization problems of this case study were formulated in GAMS 24.6.1 and solved using GUROBI 6.0 on an Intel® Core™ i7–4770 3.40 GHz with 16 GB RAM. Table 3-1 provides a breakdown of the model statistics for both models: hourly and system states. The number of constraints, nonzero elements, continuous variables, and binary variables in the hourly model are respectively 10.1, 7.7, 9.5, and 4.3 times the corresponding variables in the system states model. This reduction on the size of the problem led to lower CPU times in the system states model. The system state approximation is faster than hourly approach by a factor of 235 for 0.1% relative optimality gap, and 315 for 10^{-4}% relative optimality gap.

Table 3-1. Model Statistics for Transmission Constrained UC Case Study

Results	Hourly	System States
Constraints	108193	10759
Nonzero elements	891396	115437
Continuous variables	130537	13712
Discrete variables	25861	5951
CPU time [s] (0.1% gap)	46803	199
CPU time [s] (10^{-4}% gap)	64800[a]	215

[a] This time corresponds to the time limit, i.e., 18 hours. Therefore, the 10^{-4}% gap for this case was not reached at this time. The gap obtained at time limit was 0.0676%.

3.1.6 Results for Non-congested Network

Calculated error measurements are shown in Table 3-2 for this case. The objective function error is lower for the clustering defined using net demand per node (SS-ND) compared to the one obtained using demand and wind per node (SS-DW). However, in both situations the objective function is approximated with less than 2% error. The error measurements show the same pattern for both clustering approaches, however, SS-ND has lower values for error in 12 of the 19 error measurements compared to SS-DW. For this reason, it was concluded that using just the dimension of net demand per node to define system states resulted in a better representation of the hourly model in this case study. This situation is explained because the SS-ND has lower objective function in the clustering process (0.4506) than the SS-DW (0.5240). A smaller SED means that the system states are more closely representing a real value of a real hour and this proximity is important in the network-constrained system states model. A clustering representation with three features (demand per node, wind production per node, and net demand per node) was also analyzed, however, the distance measure at the end of the clustering process is higher than the one obtained for SS-ND. The representation of the storage level is more accurate for the long-term hydro reservoir, than it is for the BESS. Although both energy storage technologies can be used for energy

arbitrage, it is more difficult to model faster technologies such as batteries in system states approximation. That being said, the resulting storage level still followed the general shape of the hourly solution seen Figure 3-4 (top). The representation of production technologies to meet base loads such as nuclear, coal, and combined-cycle (CCGT) was more accurate than those commonly used for peak or marginal loads (i.e., open cycle gas-turbines – OCGT, and fuel oil). These peaker-plant-type generation technologies are more difficult to represent when there is no congestion on the network. In fact, for the SS-ND approach, the total amount of energy produced by gas-turbines was 5.65 times larger than the value obtained in the hourly model. As a consequence of system states approximation, errors shown in Table 3-2 associated to power flow are strongly related to errors in approximating thermal production.

Table 3-2. Error Measurements Non-Congested Network

Results	SS-DW NRMSE [%]	SS-DW AE [%]	SS-ND NRMSE [%]	SS-ND AE [%]
Objective function	N.A.	1.6	N.A.	1.6
Hydro storage level	0.2	0.2	0.2	0.2
BESS storage level	18.1	11.4	18.1	15.8
Nuclear production	1.1	0.1	0.6	0.0
Coal production	15.9	3.2	14.5	4.0
CCGT production	29.1	14.9	30.9	4.6
OCGT production	31.6	769.4	29.7	565.3
Fuel Oil production	6.6	25.3	11.4	80.7
BESS production	12.4	22.3	12.7	10.5
Line 1-2 power flow	17.8	22.2	16.6	20.2

N.A.: Not Applicable

3.1.7 Results for Congested Network

For this case study, the transmission capacity of line 1-2 is reduced to 800 MW. Table 3-3 shows the error measurements for this case study using each clustering approach. It is important to highlight that, in both clustering approaches, the objective function error is lower in the congested case. This can be attributed to the improved approximation of marginal generation technologies (i.e., OCGT and fuel oil), which had highly reduced AE values and therefore led to the model providing a better overall system cost approximation. Another observation was that although the error in measuring storage level increased, the AE in the quantity charged by the BESS decreased significantly. This improvement in calculating thermal production and charged BESS power is due to the congestion on line 1-2. The limitation of the transmission network reduces the feasible region of the optimization problem. Consequently, the optimal solution value of the hourly model and the system states approximation are closer. This increased accuracy is also observed in the reduced approximation error of the power flow through line 1-2. From the congestion network case study, it is possible to draw two main conclusions: the first one is that, despite some error measurements increases in Table 3-3 compared to Table 3-2, the overall accuracy of the system state model is improved. Secondly, as mentioned in the non-congested network case, the net demand per node (SS-ND) for clustering displayed better accuracy in the system state approach than the demand and wind per node (SS-DW) clustering

strategy. As in the non-congested case, the resulting storage level still followed the general shape of the hourly solution seen Figure 3-4 (bottom).

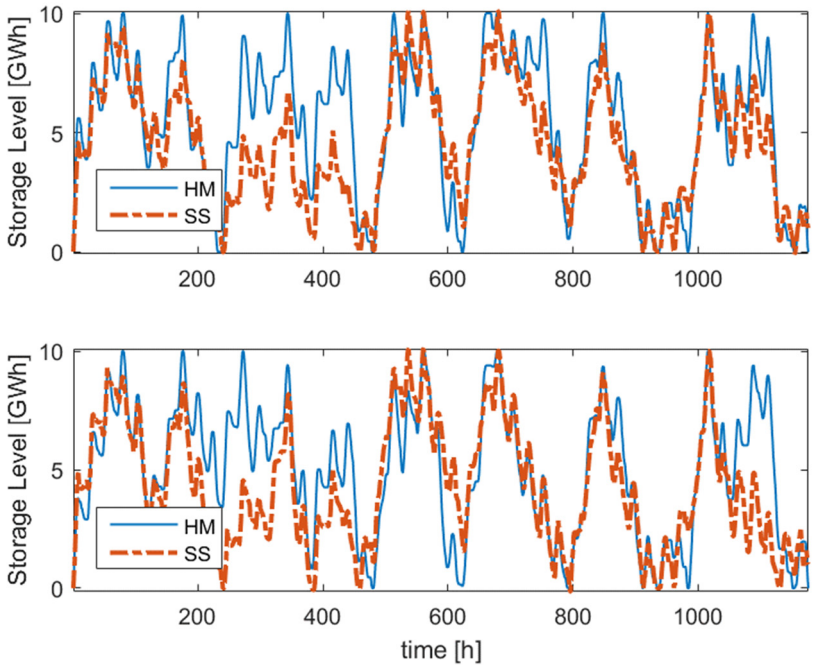


Figure 3-4. Non-congested network and congested network results

Table 3-3. Error Measurements Congested Network

Results	SS-DW NRMSE [%]	SS-DW AE [%]	SS-ND NRMSE [%]	SS-ND AE [%]
Objective function	-	1.3	-	1.1
Hydro storage level	0.2	0.2	0.2	0.3
BESS storage level	24.0	17.4	20.3	14.4
Nuclear production	5.2	1.4	0.8	0.1
Coal production	33.7	13.2	15.4	3.3
CCGT production	46.4	57.3	28.9	1.2
OCGT production	27.7	99.3	28.2	73.2
Fuel Oil production	9.8	23.6	10.1	19.8
BESS production	10.9	6.7	11.5	3.4
Line 1-2 power flow	11.8	9.8	11.6	4.4

Table 3-4 shows the weighted average energy price for the hourly model and the system states approach. Furthermore, the average prices at peak and off-peak hours are presented in order to compare the prices at extreme conditions.

Table 3-4. Weighted Energy Price [$/MWh]

Results		Average	Peak	Off-Peak
non-congested network	Hourly Model	26.27	28.08	21.87
	System State	28.97	34.89	20.33
congested network	Hourly Model	24.45	27.76	8.21
	System State	24.16	27.22	8.49

The average error in a non-congested network is near to 10%; however, with a congested network the error is reduced to 0.6%. Peak and off-peak prices

are also improved when the network is congested as a consequence of the enhancements of the overall accuracy of the system state model. Figure 3-5 shows the weighted energy price in both models; for the sake of simplicity only the first 500 hours are plotted, the rest of the hours showed a similar behavior.

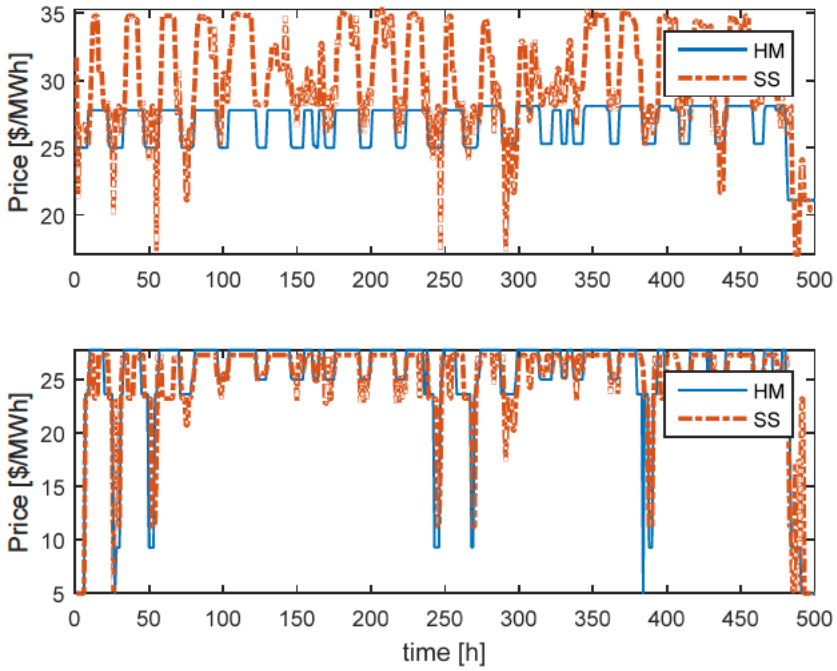


Figure 3-5. Non-congested network and congested network results

3.1.8 Sensitivity Analysis

Three sensitivity analyses are conducted over the case study shown before. First a sensitivity analysis related to the number of system states. Second, an analysis exploring different levels of congestion in the line 1-2. Finally, a sensitivity analysis surveying the effects of simultaneously congesting multiple elements in the network (i.e., lines and transformers).

1. *Sensitivity to number of system states*: As mentioned before, system states are obtained through a clustering technique. Therefore, the number of clusters is a key issue in applying this framework. Table 3-5 shows for different number of clusters or system states the distance measure (i.e., squared Euclidean distance - SED), objective function error with respect to the hourly model, and CPU time. It can be observed that adding further clusters beyond 100 does not have a significant effect on the model since the errors in the objective function, charged power, and storage level in the BESS remains relatively stable. In fact, the difference in the objective function error between 100 clusters and 200 clusters is only 0.24%. Another important conclusion to be drawn is that generally, adding more clusters increases CPU time. This result makes sense in that additional clusters leads not only to growth in the number of system states but also to the size of the frequency matrix (see section 2.3.1 and appendix A.3). It is important to note that there can be exceptions for specific cases since we are solving MIP models. For example, for 80 clusters the CPU time is longer than the one value obtained with 100 clusters. This behavior is an exception and is a result of using a MIP model. The CPU time will depend

also on the Branch and Bound algorithm applied by the solver. For this particular case study, 100 clusters (8.53% of the 1172 hours in the hourly UC of reference) displayed a good balance between approximation error and solution time.

Table 3-5. Sensitivity to number of clusters or system states

SS Number	Frequency matrix size	SED	CPU time [s]	Obj. Func error [%]	BESS production NRMSE [%]	BESS storage level AE [%]
60	11394	0.827422	41	1.7	13.2	16.2
80	14486	0.595395	495	1.6	13.4	14.5
100	16471	0.450577	215	1.6	12.7	15.8
120	18234	0.352573	1803	1.4	12.4	14.2
200	23759	0.173679	14005	1.3	12.8	15.6

2. *Sensitivity to different levels of congestion in line 1-2*: In Sections 3.1.6 and 0, the results for two case studies—with and without congestion in line 1-2—were presented. We observe that congestion in the network improves the system state model's accuracy. To check the validity of this conclusion, different levels of congestion along the line 1-2 were analyzed. Four different congestion levels were analyzed: *none* (line 1-2 limit =1500 MW), *low* (line 1-2 limit =800 MW), *moderate* (line 1-2 limit =750 MW), and *high* (line 1-2 limit =700 MW). These levels correspond to 0%, 22%, 48%, and 89% congestion of the time horizon respectively. Table 3-6 compares, the errors obtained for the objective function, charged power, storage level in the BESS, and power flow in line 1-2 for each congestion level. There is an observed decreasing trend in the errors amongst these criteria as congestion levels increase. The main reason for this behavior is the fact that a congested network limits the feasible region of the optimization problem in both models. Therefore, the set of possible values for each variable is smaller than in the case without congestion. In fact, the power flow in line 1-2 is the variable that improves most when the congestion level increases. It is important to note that the objective function error for moderate congestion level is a slightly lower (0.02%) than the value for high congestion level. However, both values are lower than the objective function for the non-congestion level. Consequently, it is still possible to conclude for this particular case that if the network is congested then the approximation of the objective function is improved.

Table 3-6. Sensitivity to congestion level

Congestion levels	Obj. Function error [%]	BESS production NRMSE [%]	BESS storage level AE [%]	Line 1-2 power flow NRMSE [%]
None (1500 MW)	1.6	12.7	15.8	16.6
Low (800 MW)	1.1	11.5	14.4	11.6
Moderate (750 MW)	1.1	9.6	13.7	7.8
High (700 MW)	1.1	8.9	11.8	3.0

3. *Sensitivity to multiple elements congested*: In the previous section, the results for congestion in line 1-2 were analyzed. In order to identify the behavior of the system states approach when there is more than one congested element, different scenarios with reduced transmission capacity in transformers 5-6, 4-9, and 4-9-8 were analyzed (transformer 5-6 limit= transformer 4-9 limit = transformer 4-9-8 limit = 400 MW). Table 3-7 shows the error metrics for these cases. As in the case study for congestion only in line 1-2, it is possible to observe that error measurements are improved for marginal thermal generation technologies, BESS charge, and power flow in line 1-2. It is important to note that the objective function error is even lower in this case than in the case study with only one congested element in the network. This is again related to the reduction in the feasible region due to the congestion in the network.

Table 3-7. Sensitivity to multiple elements congested

Results	SS-ND NRMSE [%]	SS-ND AE [%]
Objective function	-	0.7
Hydro storage level	0.3	0.3
BESS storage level	10.1	6.9
Nuclear production	0.6	0.0
Coal production	18.0	4.5
CCGT production	39.8	25.2
OCGT production	8.9	85.2
Fuel Oil production	19.8	64.4
BESS production	10.9	2.8
Line 1-2 power flow	13.3	0.1

3.2 System States and Representative Periods Comparison

This section compares the system states and representative periods methods that we have described in Section 2.3. Both short-term energy storage (STESS) and seasonal storage (LTESS) are included in the optimization models in order to analyze different types of ESS technologies. In addition, enhanced versions of each method are proposed in order to overcome the main drawbacks in their former methods. A Spanish case study is used to test all methods and draw conclusions at the end of this section.

3.2.1 Optimization Models

Five optimization models are solved in this section. All optimization models are detailed in Appendix A.2 to A.6:

- Hourly UC Model (HM), which is used as a benchmark.
- The classic System States (SS).
- The System States Reduced Frequency Matrix (SS-RFM), which is the enhanced version of the SS model. Section 3.2.3 explains the main difference between this model and the classic SS model.
- The classic Representative Periods (RP).

- The Representative Period with Transition Matrix and Cluster Indices (RP-TM&CI[11]), which is the enhanced version of the RP model. Section 3.2.4 explains the main difference between this model and the classic RP model.

Figure 3-1 shows an overview of the analysis. We have as an input data the hourly demand, wind, and solar time series. Therefore, we can solve the HM model to obtain the benchmark results. In addition, two different clustering procedures are applied. First, we cluster individual hours in order to obtain the system states for the SS and SS-RFM models. Second, we cluster grouping by days the time series in order to obtain the representative periods for the RP and the RP-TM&CI. Finally, the hourly results of each model are compared to determine the quality of the approximations in terms of objective function cost, productions, stored energy, and prices.

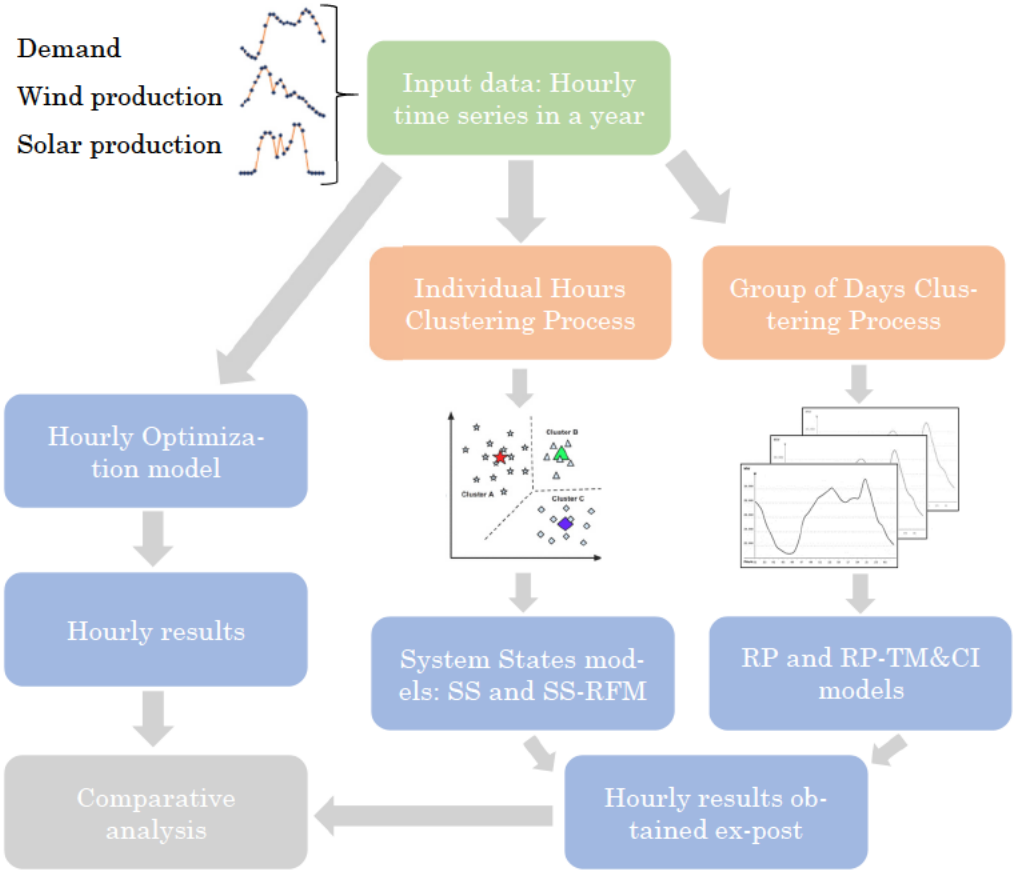


Figure 3-6. Analysis overview: comparison of SS and RP models

3.2.2 Main Drawbacks of SS and RP Models

The SS and RP models have some drawbacks, which are detailed in a case study in Section 3.2.6. In this section, we summarize these drawbacks:

- The SS model results and CPU time are highly dependent on the size of the *frequency matrix* (see Appendix A.3). This matrix is used to guarantee that storage levels are between the maximum and minimum bounds for each storage unit throughout the time horizon. Furthermore, it helps to

[11] In Chapter 4, the RP-TM&CI model is renamed as the Linked Representative Periods (LRP), because this name fits better in the context of that Chapter.

keep some chronological information for ESS in the optimization process. However, this proposal has two disadvantages. First, short-term storage devices such as batteries require several bounds in a day to ensure that storage levels are within bounds, but the greater the number of bounds, the greater the frequency matrix size and, therefore, the longer the CPU time. Second, in order to determine the number of bounds we need an iterative process called Iterative Bounding Method (detailed in Appendix A.3), which adds even more CPU time to the SS model.

- The RP model solves each representative period (e.g., day) independently and with the same constraints as the HM model. CPU time thus depends on the number of representative periods instead of on the number of bounds for storage units, as it does in the SS Model. The main drawback is that chronology among the representative periods is lost and storage levels of storage units with a cycle longer than the representative period (e.g., hydro units) are not determined adequately. This is especially important in hydrothermal power systems or power systems with pumped hydro storage potential.

In the following sections, we propose enhanced versions of the SS and RP models to tackle these drawbacks. These models were published in [95].

3.2.3 Enhanced System States Model

The main drawback in the classic system states modeling is the increase of computational burden when several storage bounds constraints are needed. Therefore, the System States Reduced Frequency Matrix Model (SS-RFM) is proposed in this thesis to overcome this drawback. Instead of using the *frequency matrix* ($F_{s,s',k}$) to establish the storage bounds constraints, the SS-RFM uses a *reduced frequency matrix* ($RFM_{s,s',k}$) in these constraints. Figure 3-7 shows an example to explain the difference between these two matrices. In this example, we have a set of twelve hours $h = \{\text{h1}, \text{h2}...\text{h12}\}$, four system states $s = \{\text{s1}, \text{s2}, \text{s3}, \text{s4}\}$, and three hours in which we want to impose storage bound constraints $k = \{\text{h4}, \text{h8}, \text{h12}\}$. The example shows the *cluster index*, which is the vector that contains the relationship between hours and system states. For instance, the hours h1, h3, h4, h7, and h9 belong to system state s1, while h12 is the only hour that belongs to system state s4. The frequency matrix is calculated for each hour in the set k as the total number of transitions among the system states considering the cluster index vector. In Figure 3-7, the frequency matrix at hour $k = \text{h4}$ has three values that correspond to the transitions among system states from h1 to h4. The reader can verify that the same logic applies to the frequency matrix at hours h8 and h12. Here it is also possible to observe that the frequency matrix increases (as well as the non-zero elements in the optimization problem) every time we want to impose storage bounds constraints. This is the main reason why the computational burden increases when several hours are considered in the set k.

The Reduced Frequency Matrix ($RFM_{s,s',k}$) is just the difference between the frequency matrix corresponding to the current hour k and that corresponding to the previous element in set k, that is, $k - 1$. In other words, the difference between these two elements or hours in the set k could be understood as a moving window. In Figure 3-7, the RFM at h8 is the difference between the frequency matrix at h8 and h4, and the RFM at h12 is the difference between

the frequency matrix at h12 and h8. The RFM at each hour in set k remains with the same number of elements if the moving window is constant. Therefore, the non-zero elements in the optimization problem do not increase with the number of storage bounds constraints. This leads to an advantage in the CPU time that is shown in Section 3.2.6. However, it is important to mention that despite the use of the RFM, the storage level could be out of bounds because the hours in set k are predefined in the model and we do not know in advance the storage level value at each hour in set k. The best practice for reducing the number of hours in which the storage levels can be out of bounds is to predefine the moving window considering the smallest storage cycle in the power system (e.g., 4h to 8h for a BESS).

The complete optimization problem for the SS-RFM is shown in Appendix A.4.

Cluster Index

hours →	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10	h11	h12
system states →	s1	s2	s1	s1	s3	s2	s1	s2	s1	s3	s3	s4

chosen hours to impose bounds in the SS models: $k = \{h4, h8, h12\}$

Frequency Matrix $F_{s,s',k}$

k=h4	s1	s2	s3	s4
s1	1	1	0	0
s2	1	0	0	0
s3	0	0	0	0
s4	0	0	0	0

k=h8	s1	s2	s3	s4
s1	1	2	1	0
s2	2	0	0	0
s3	0	1	0	0
s4	0	0	0	0

k=h12	s1	s2	s3	s4
s1	1	2	2	0
s2	3	0	0	0
s3	0	1	1	1
s4	0	0	0	0

Reduced Frequency Matrix $RFM_{s,s',k}$

k=h4	s1	s2	s3	s4
s1	1	1	0	0
s2	1	0	0	0
s3	0	0	0	0
s4	0	0	0	0

k=h8	s1	s2	s3	s4
s1	0	1	1	0
s2	1	0	0	0
s3	0	1	0	0
s4	0	0	0	0

k=h12	s1	s2	s3	s4
s1	0	0	1	0
s2	1	0	0	0
s3	0	0	1	1
s4	0	0	0	0

Figure 3-7. Frequency matrix and reduced frequency matrix

3.2.4 Enhanced Representative Periods Model

The Representative Period with Transition Matrix and Cluster Indices (RP-TM&CI) model is the second original contribution of this thesis. This model overcomes the main drawbacks of the classic representative periods model by introducing two main concepts: 1) *linking representative periods using a transition matrix*, 2) *superposing inter- and intra-period storage balance constraints*. Both concepts are explained through the following examples in Figure 3-8 and Figure 3-9. Although the RP-TM&CI model is sufficiently general to be able to work with representative periods of any length, we will speak of representative days for the sake of simplicity in these examples.

Figure 3-8 shows how the representative days are linked. In this example, we have a set of seven days $d = \{d1, d2, ..., d7\}$ (e.g., seven days in a week) and a set of two representative days $rp = \{rp1, rp2\}$ (e.g., one day to represent the working days and other to represent the weekend days). We show the *cluster index*, which is the vector that contains the relationship between days and representative days. For instance, the days d1 to d5 belong to rp1, while d6

and d7 belong to rp2. The *transition matrix* is calculated using this information: four times we jump from rp1 to rp1, one time from rp1 to rp2, and one time from rp2 to rp2. This information is used to link the days and create continuity between the representative days. Therefore, the thermal units that are committed in the last hour of the first RP are also committed in the first hour of the second RP, see equation (A.6-1) in Appendix A.6. As a consequence in Figure 3-8 example, the thermal units that are on in the last hour of rp1 are also on in the first hour of rp2 as well as in the first hour of rp1. Notice that last hour of rp2 is not linked to the first hour of rp1 since the transition from rp2 to rp1 is zero in the transition matrix. As written here, if there is even one transition between the two days, this constraint is applied. However, the constraint could be set to take effect only if there is a considerable number of transitions between the two days, 5 or 10% of the transitions in the time horizon, for example. This consideration in the current example would lead to only linking last hour of rp1 to its first hour since the other transitions would be disregarded.

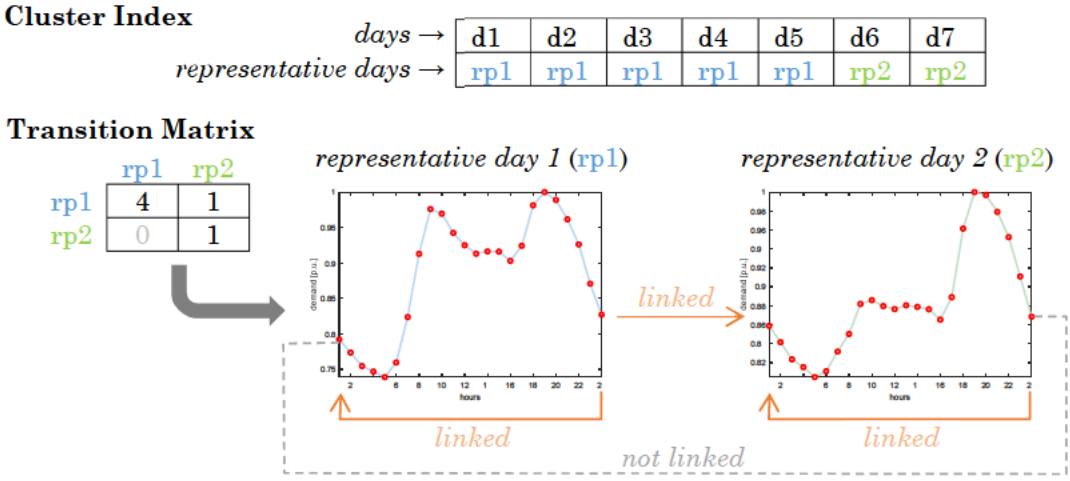


Figure 3-8. Example of transition matrix use in RP method

Figure 3-9 shows the second innovation of this model; it creates the continuity in storage across the entire time horizon that allows for the modeling of long-term storage. Once more, we take advantage of the information in the *cluster index*. Since we have the hourly chronology within each representative day, it is possible to determine the hourly storage level and, therefore, imposing bounds to this variable. These constraints are called *intra-day* (or intra-period) storage balance equations. In this example, we obtain 24 equations per ESS for each representative day. These constraints are imposed only to short-term energy storage (STESS) because they are more likely to do more than a full cycle within the representative period (e.g., a BESS with a 4-hours discharge time could have more than one full charge/discharge cycle in one day). We also propose a set of constraints that checks at regular intervals (e.g., 1 week) the ESS storage level. These constraints are called *inter-day* (or inter-period) storage balance equations, and they are defined using the *cluster index* information. In this example, the storage level at the end of the week (i.e., hour 168) is the storage level at the beginning (i.e., initial storage level) plus/minus 5 times the sum of all charge/discharge decisions in rp1 and plus/minus 2 times the sum of all charge/discharge decisions in rp2. The rp1

decision variables are multiplied by 5 because, according to the cluster index vector, the rp1 represents days d1 to d5. The same logic applies to rp2, which represents 2 days within the example. In other words, this model is superposing the inter- and intra-period storage balance constraints to create chronological continuity in the problem. These is highly important to model seasonal storage and reducing the time representation at the same time. The inter-day balance equations are imposed to both STESS and LTESS because it allows both ESS to internalize long-term signal in the short-term decisions. In Chapter 4, we analyze this situation using the storage/water value concept in the RP-TM&CI model.

The complete optimization problem for the RP-TM&CI is shown in Appendix A.

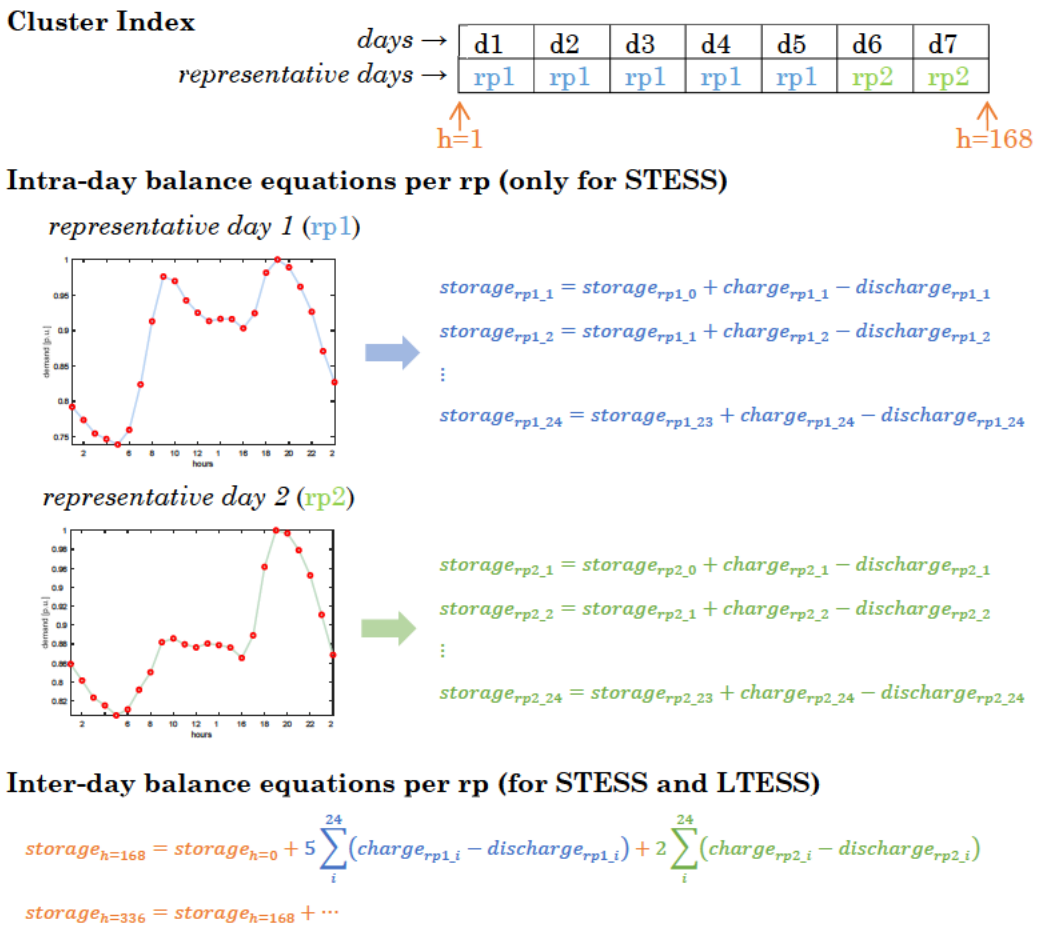


Figure 3-9. Example of superposing inter- and intra-day in ESS balance

3.2.5 Case Study Description

As a case study, we chose the Spanish power system in target year 2030. The Spanish case is interesting because it has hydro reservoirs (i.e., LTESS with monthly or yearly cycle) and, according to ENTSO-E [100], the next ten years will likely bring investment in Battery Energy Storage System (BESS) and Pumped Hydroelectric Energy Storage (PHES), i.e., STESS with daily or weekly cycle. We ran four different scenarios for 2030 on the hourly model and the four approximate models. The wind and solar profiles for these scenarios were taken from [101], [102] while hourly demand data and annual production per technology were taken from the ENTSO-E *Ten Year Network*

Development Plan 2016 [100]. Scenario 1 and 3 were based on national predictions, whereas scenarios 2 and 4 were designed with the whole of Europe and climate protection goals in mind. The scenarios include a significant development of renewable electricity sources, supplying 35% to 60% of the total annual demand, depending on the Scenario. Moreover, the hourly demand curve of each Scenario reflects the potential for demand response, which rises from 5% in Scenario 1 to 20% in Scenario 4.

For each of the four scenarios, the SS and RP models were run with four different numbers of clusters for increasing time resolution. The RP and RP-TM&CI models used 4, 9, 18, and 37 representative days which corresponds respectively to 1%, 2%, 5% and 10% of the time horizon. Time resolution within each representative day is hourly. The SS and SS-RFM models used 26, 48, 96, and 216 system states. These numbers of states were chosen because they provided a *fair* comparison with the clusters used with the RP models by having roughly the same number of binary variables.

The representative days were chosen by normalizing time series for the hourly demand, wind availability, solar availability, and hydro inflows, and combining 24 hours of those time series (96 dimensions in all) into a single point to be clustered with the rest of days of the year using k-medoids. The system states were chosen in an analogous manner. The four-time series were normalized, but this time each point to be clustered represented only one hour (4 dimensions) and the clustering method was k-means so that the resulting system state was the centroid of the cluster (a composite hour) rather than a true hour.

3.2.6 Case Study Results

For this case study, we considered a total BESS installed capacity of 10 GWh with a maximum output of 1GW and a 0.9 efficiency coefficient. Figure 3-10 shows a box & whisker plot for CPU Time while Figure 3-11 objective function error considering the results for each scenario. All models were solved until optimality, i.e., until the integrality gap equaled 10^{-6}. Figure 3-10 shows the time necessary for the solution of each model as a fraction of the time taken by the hourly model as the number of clusters (i.e., system states or representative days) increases.

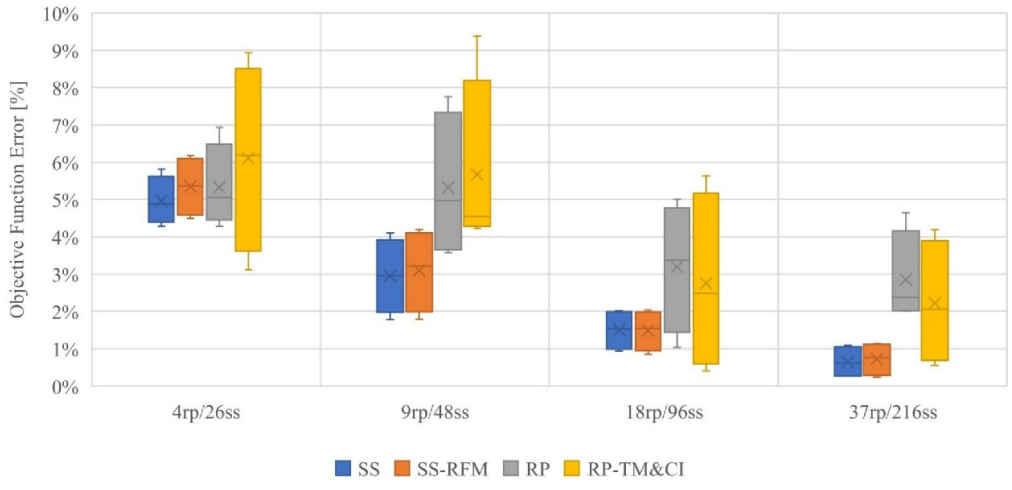


Figure 3-10. CPU time for each tested model

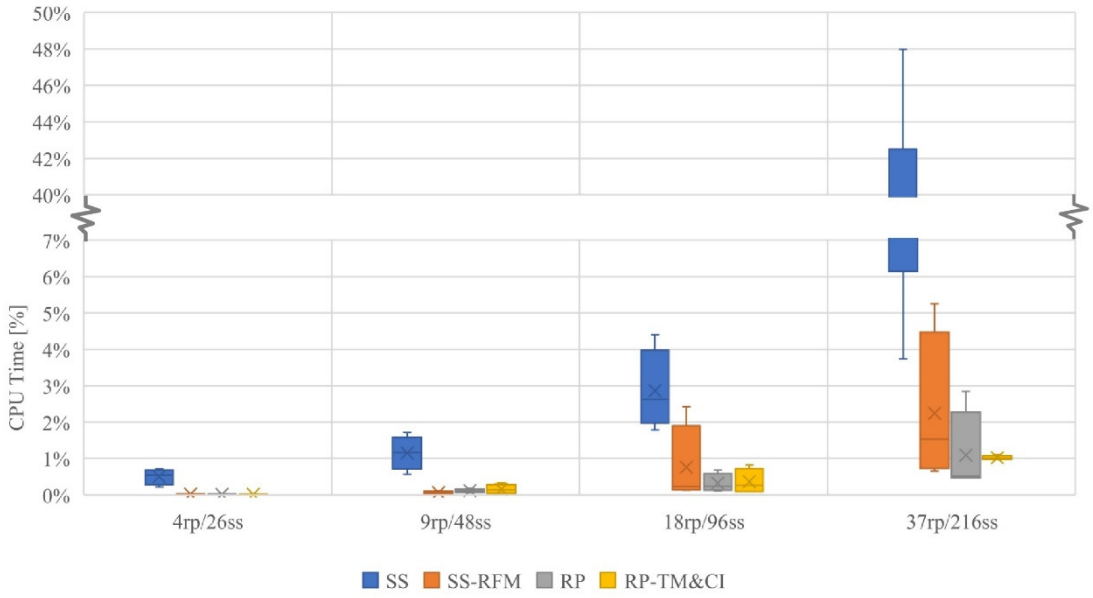


Figure 3-11. Objective function error for each tested model

As expected, the amount of time necessary for model solution increases with the temporal resolution, but up to the 3rd time resolution (18rp, 96ss) all four approximate models took less than 5% of the time that the hourly models took. Moreover, as expected, increasing the number of system states or representative days reduced the error in the objective function, see Figure 3-11. In addition, the results show the improvement obtained with the SS-RFM and RP-TM&CI models proposed in this thesis. The SS-RFM model took between 4 and 20 times less CPU time than the SS model without hampering the performance of the approximation in the objective function error. Moreover, the RP-TM&CI model reduced the objective function error of RP model as the number of representative days increase without a significant rise in the CPU time. For the sake of simplicity, the rest of this section shows only the results for the 3rd time resolution (18rp, 96ss) because it has a good trade-off between CPU time and objective function error.

Table 3-8. Average errors [%] for each tested model

	Result	SS	SS-RFM	RP	RP-TM&CI
Production	Nuclear	-0.3	-0.2	5.4	-0.2
	Coal	1.9	1.2	10.5	-2.0
	CCGT	2.3	2.8	-10.6	1.3
	Hydro	-0.2	-0.2	-10.4	0.8
	BESS	7.3	11.3	-17.0	-4.8
	Renewable	-0.5	-0.5	-0.4	-0.5
VRES curtailment		24.7	24.9	18.4	18.6
Start-up	Coal	-53.9	-54.3	-52.4	-9.3
	CCGT	-73.6	-75.2	-91.3	-21.0
Price	Average	-0.5	0.03	8.0	0.7
	Max	-25.4	-8.5	-22.7	2.1
	Min	0.0	0.0	0.0	0.0

So far, we have used objective function error to judge the accuracy of the approximate models, nevertheless, results such as annual production per technology, total number of startups, and energy prices allow for a more detailed comparison. Table 3-8 shows the average error for these results when comparing each approximate model to the hourly model. Negative values in Table 3-8 show overestimation in the approximate model while positive values are underestimation. For thermal production SS, SS-RFM, and RP-TM&CI models have errors lower than 3% while the RP model has error between 5% and 11% because it solves each representative day individually. The SS and SS-RFM models give the estimation of total hydro production closest to that of the hourly model while the RP model gives a very poor estimate. This is because the RP model constrains the storage at the end of each day to be higher than at the beginning so hydro storage cannot evolve according to its natural yearly cycle. The RP-TM&CI model, however, does succeed in estimating the annual hydro production, which is what it was designed to do. The SS and SS-RFM models do not approximate the annual BESS production very well, as the models cannot keep the energy fully within bounds throughout the time horizon. The RP-TM&CI model gives a value of the total annual BESS production that is closest to the HM model. VRES production is estimated with good accuracy (i.e., errors less than 0.5%) for all models, while the VRES curtailment has more error and is underestimated in all models. However, representative periods-type models have slightly better accuracy than system-states-type models. The RP model overestimates the number of necessary startups during the year of peaking units (CCGT), which is only to be expected since it treats each day as separate from the others. Because they maintain some chronology between periods using the transition matrix, SS and SS-RFM do a better job of estimating startups than the RP. However, the RP-TM&CI model has the number of startups closest to that of the HM model, as it uses its transition matrix to keep continuity between the thermal units at the end of one day and the beginning of the next. These results also demonstrate the effectiveness of the RP-TM&CI model over the RP model. In the case of the energy prices, the RP model makes the worst estimate due to the previous results. The average prices in SS, SS-RFM, and RP-TM&CI are all quite accurate, but the maximum price is better estimated in the enhanced models, SS-RFM and RP-TM&CI. This is important because the storage investment results are partially driven by the differences between the maximum and mini-mum prices. We analyze this situation in Chapter 5.

Figure 3-12 shows the storage level evolution for hydro unit in scenario 1, while Figure 3-13 shows it for the BESS. Not only is the total yearly hydro production estimated by SS, SS-RFM, and RP-TM&CI very close to that of the HM as shown in Table 3-8, but the overall storage evolution closely follows that of the HM, Figure 3-12. The RP model cannot correctly estimate the evolution of storage levels considering the production, consumption, inflows, and spillages for each representative day because the representative days are not related among themselves. The RP-TM&CI model fixes this by considering chronology among the representative days using the transition matrix and cluster indices. In fact, the RP-TM&CI model yields the prediction of hydro storage levels that is most similar to that of the HM model.

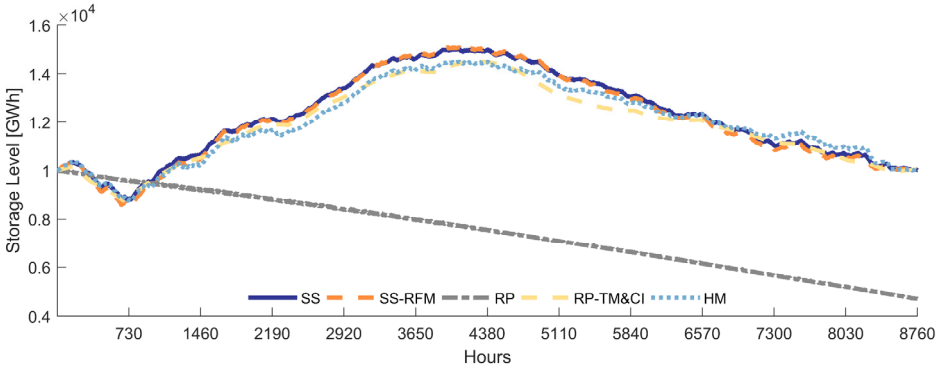


Figure 3-12. Hydro storage level for each tested model

The BESS storage level is shown in Figure 3-13 for a week of the year. RP and RP-TM&CI models perform best when the BESS charge and discharge in a single day. If, however, the true BESS charges and discharges over the course of more than one day then the RP and RP-TM&CI have trouble approximating that, as they are limited to the representative days. Despite this, the RP-TM&CI model performs better than the RP model due to the chronological information shared among the representative days. The SS and SS-RFM models have better performance than the representative days models because they are not limited to the period length, i.e., 24 hours, and this allows them to capture charging and discharging periods longer than a day. However, the SSs models cannot guarantee that BESS storage levels stay within bounds.

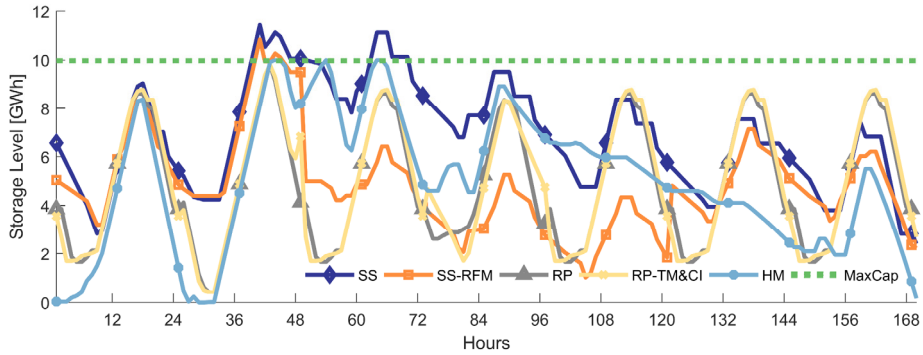


Figure 3-13. BESS storage level for each tested model

In Figure 3-13 both SS and SS-RFM predict that BESS storage levels will exceed the upper bound, which is unrealistic in a power system operation. To correct that behavior, the number of constraints should be increased, but this vastly increases CPU time in the SS model and increases the error in the SS-RFM model. If the extra constrained hours are chosen using the iterative method, this increases the CPU time still further.

3.2.7 Discussion

In this section we want to highlight the link between short- and long-term storage. In this thesis we focus on modeling energy storage operation operational details, considering long-term (i.e., seasonal) hydro storage generation as well as short-term (i.e., hours) storage systems such as batteries. These are very different resources in the power system. Therefore, the following question arises: why try to model both with the same methodology? To answer this

question, we need to consider that hydro storage already exists in most real power systems and more could be built in the future, and short-term storage (e.g., BESS) is getting cheaper and could be a good technical solution to reduce VRES curtailments even with relatively low energy to power ratios (e.g., 1-4 h). Moreover, if both types of storage are not considered at the same time, then an assumption must be included regarding storage operation. For example, it is possible to consider maximum available hydro energy without tracking the storage level, or to assume a peak shaving for short-term ESS. In either case, one decision is fixed while the other is optimized. Therefore, possible synergies between both storage systems are neglected. This is the case of more traditional hydrothermal dispatch models.

The RP-TM&CI model co-optimizes both types of storage. Hence, the operational decisions of short- and long-term storage are now linked and depend on each other. The benefits of this co-optimization are shown in the results. In fact, the best results are obtained with the RP-TM&CI model, which represents the relationship between both types of storage better than the other approximate models. It should also be noted that the RP-TM&CI model could be used to improve traditional hydrothermal models in which the water value serves as a consistent way of coupling long-term reservoir management with short-term operations of storage units. Using the RP-TM&CI model it might be possible to obtain the water value of long-term reservoirs internalizing the information of short-term storage, which is not possible in traditional hydrothermal models. Chapter 4 analyzes the application of the RP-TM&CI model in the hydrothermal dispatch problem.

3.3 Main Takeaways in this Chapter

Through this chapter we have analyzed several optimization models for ESS operation in medium- and long-term planning. These analyses are divided in two studies. First, we have evaluated the impact of transmission network in the clustering methods used to reduce the temporal information while conserving chronological information for the ESS. Second, we have compared the classic versions of system states and representative periods methods when short-term energy storage and seasonal storage are included in the optimization models. In addition, enhanced versions of each method have been proposed in order to overcome the main drawbacks in the classic methods.

The main analysis and results in this chapter can be summarized in the following conclusions:

- *The system states approach can be extended to include transmission network constraints, it requires an enhancement of the method used for clustering the hourly time periods. It is shown that this extension does not hamper its computational benefits.*
- *If the transmission network is constrained, the accuracy of the system state approach improves. This conclusion is quite relevant and somehow counterintuitive, because it means that it is not necessary to include in-formation of the congestion of the network in the clustering-process to improve the results when the system state framework is used to represent ESS operation in MLTOP models.*

- *The SS model was originally developed to include chronology and high time resolution details in MLTOP models. While it can deal with long-term storage (LTESS), it cannot accurately estimate short-term storage (STESS), and quickly becomes computationally intensive because of the storage constraints. The SS-RFM model takes much less time to run than the regular SS model, because it reformulates the storage constraints, but it does not improve the accuracy of the short-term storage modeling. Moreover, SS models could lead to infeasible results (i.e., more energy stored than the maximum storage capacity), which is their major drawback, and means that they require additional adjustments for most practical applications.*
- *Unlike the SS models, the RP model cannot handle long-term storage (LTESS), however, it deals well with short-term storage (STESS) as it preserves within-day chronology. The RP-TM&CI model combines aspects of the SS and RP models to account for both short and long-term storage. According to the case study results, it is the most accurate of the four approximate models and does not require a significant increase of CPU time.*
- *We expect that this behavior can be replicated in different case studies, because the results in this chapter support the idea that including chronological information among representative periods may be an efficient way to include small time scale variations in longer-term planning models that involve storage. Doing so is a critical need in the adequate representation of power systems that include a significant and increasing quota of variable renewable sources and energy storage systems.*

4 HYDROTHERMAL DISPATCH USING LINKED RP

Contents

This chapter proposes a new formulation to include STESS operational decisions in a stochastic hydrothermal dispatch model, which is based on the proposed enhanced version of the representative periods approach that allows an analysis of both short- and long-term storage at the same time. For the analyzed case study, this approach models operating decisions of STESS with errors between 5% to 10%, while the classic Load Duration Curve (LDC) approach fails by an error greater than 100%. Moreover, the LDC model cannot determine opportunity costs on an hourly basis and underestimates the water value by 6% to 24% for seasonal hydro reservoirs. On the other hand, the proposed LRP model produces an error on the water value lower than 3% and can determine hourly opportunity costs for STESS using dual variables from both intra- and inter-period storage balance equations. Therefore, hourly opportunity costs in the LRP model successfully internalize long-term signals due to seasonality in hydro reservoirs.

The analysis and results in this chapter were published in Article IV [103] *as working paper, which is under review at the time of this thesis publication.*

4.1 Current Hydrothermal Dispatch Models

Among the hydrothermal commercial tools utilized in scientific research, the main ones are: SDDP developed by PSR, PLEXOS Integrated Energy Model by Energy Exemplar, ProdRisk by SINTEF, NEWAVE by CEPEL, and Star-Net Model by IIT. For medium- or long-term studies, these tools use a Load Duration Curve (LDC) approach (a.k.a., load-levels approach) with monthly or weekly stages. This is mainly due to the computational efficiency of LDC for large-scale systems. However, the LDC approach lacks chronological information within stages (e.g., weeks or months) and fails to represent short-

term constraints (e.g., ramps, storage balance, etc.) [48]. Despite this situation, no relevant work has aimed to improve the representation of short-term operational decisions in hydrothermal dispatch models.

As stated in Section 2.3.2, the *representative periods* (RP) method has been applied to long-term models in order to consider either renewable energy variability in the short-term [52] or UC constraints [104]. Generation dispatch and investment decisions are made for the selected periods (e.g., days or weeks). The RPs preserve the internal chronology of their hours, rendering a more realistic representation of changing storage levels over the course of a day or week. However, the basic definition of the RP does not preserve the chronology among them. Therefore, any Energy Storage System (ESS) with a full charge-discharge cycle longer than the RP (e.g., monthly or yearly) will not be adequately represented. In order to improve this situation, in Section 3.2.4, the proposed model RP-TM&CI overcomes this shortcoming [95]. The continuity among the RPs is included through the superposition of intra-period and inter-period storage balance equations. This idea is also analyzed by some authors in [57], [105]. In this chapter, we rename the RP-TM&CI as Linked Representative Periods (LRP) in order to emphasize in the concept of representative periods that share information among them. Despite these developments in short-term and seasonal storage interaction, opportunity costs in stochastic modeling have not been analyzed. Opportunity costs for energy storage systems are not as intuitive as in the LDC models due to the superposition of both balance equations, i.e., intra-period and inter-period. Therefore, in Sections 4.3 and 4.4, the LRP model is extended to a stochastic hydrothermal dispatch model to define the opportunity costs for energy storage, considering short- (intra-period) and long-term (inter-period) operation.

The selection of RPs is an important aspect of the RP approach. Some authors have debated about the optimal length for RPs [106]. For instance, in [58], the authors suggested representative groups of days or representative weeks, which gives the advantage of increasing the amount of chronology preserved. However, the effectiveness of linking shorter RPs versus longer RPs has not been analyzed in the LRP approach. In Section 4.5, we analyze the impact and draw conclusions regarding our proposed model.

The challenge that we have tackled in this chapter is to obtain the hourly opportunity cost for storage technologies that usually operate on very different time scales. For example, batteries might have a full charge/discharge cycle within a couple of hours or days, whereas a pumped hydro storage facility – depending on the size of the reservoir – could have cycles of weeks, months or even years. Other important aspects of hydrothermal dispatch models such as uncertainty modeling [107] are out of its scope. A general formulation is proposed based on stochastic programming, which is compatible with different techniques to solve the dimensionality problem such as scenario reduction [108] and stochastic dual dynamic programming [109], [110].

In this context, the main contribution of this chapter is the derivation and analysis of the hourly opportunity cost of storage technologies when the proposed LRP model is solved. In other words, the LRP model can obtain an approximation of the ESS hourly opportunity cost within the studied time horizon without solving an hourly model. Moreover, the LRP model gets hourly

opportunity cost for different types of ESS technologies which operate on different time scales (hydro vs battery). So far this has not been possible because classic LDC-type models lack chronological information among the load levels.

4.2 Hydrothermal Topology and Scenario Tree

In the context of hydrothermal dispatch, it is important to establish the hydro topology in order to determine the relationship among the water basins. This is important to determine the *opportunity cost* or *water value* in hydro generators. Figure 4-1 shows an example a hydro topology where three reservoirs (r1, r2, r3), including their hydro units (h1, h2, h3), are related among them. For instance, reservoir r3 receives its hydro inflows, turbined water and spilled water from reservoirs r1 and r2, and the pumped water from its own hydro unit h3.

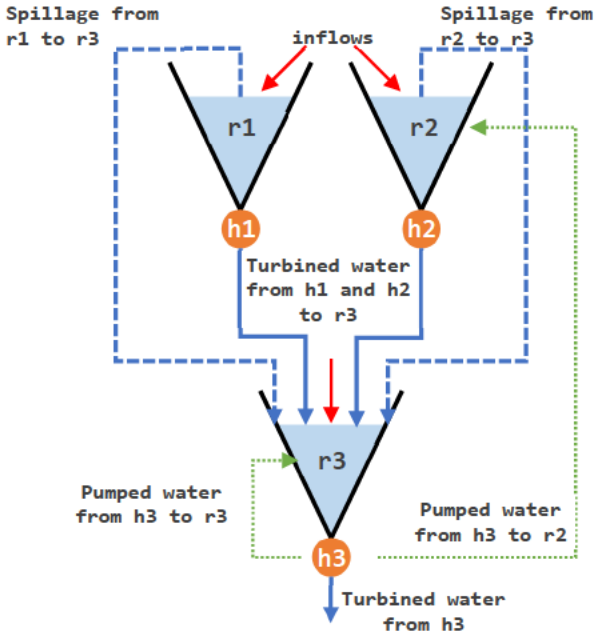


Figure 4-1. Example of hydro topology or water basin

The main source of uncertainty in hydrothermal power systems is the water inflows [111]. This uncertainty is normally represented as a scenario tree [107], in which each node in the tree represents a hydro inflow level with a certain probability. In addition, the nodes are related among them, creating different scenarios to sample different realizations of the hydro inflows. Figure 4-2 shows an example of a simple scenario tree with three scenarios: wet season, average inflows, dry season. In this example, the first stage decisions are taken in the first month represented in the tree (i.e., October) and the second stage decisions are taken for the following months. In addition, for each month in the second stage there are three different value of hydro inflows from each scenario.

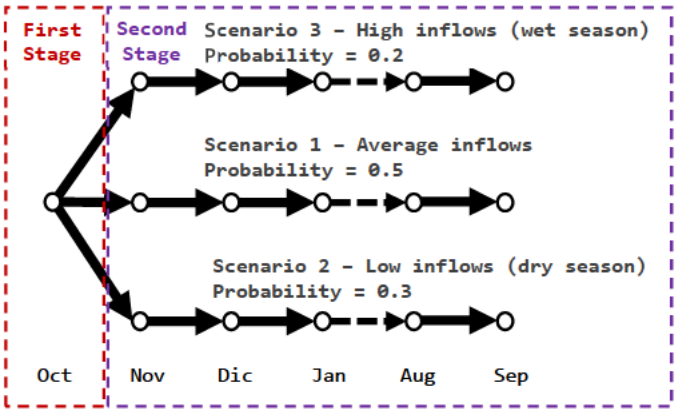


Figure 4-2. Scenario tree example

Both hydro topology and scenario tree shown in this section are used as a reference in the remainder of this chapter.

4.3 Optimization Models and Time Division

Three optimization models are solved in this chapter using a stochastic formulation. These models are detailed in Appendix B.2 to B.4:

- Hourly Model (HM), which is used as a benchmark.
- The classic Load Duration Curve (LDC).
- The Linked Representative Periods (LRP), which is the stochastic extended version of the Representative Period with Transition Matrix and Cluster Indices (RP-TM&CI) proposed in Section 3.2.4.

Figure 4-3 shows an overview of the analysis in this section. We have as an input data the hourly demand, wind, and solar time series. Therefore, we can solve the HM model to obtain the benchmark results. In addition, two different clustering procedures are applied per each node of hydro inflow uncertainty, (i.e., per month). First, we cluster individual hours in order to obtain the load levels (as in the system states approach) for the LDC model per month. Second, we cluster grouping by periods (e.g., days) the time series in order to obtain the representative periods for the LRP model per month. Finally, the hourly results of each model are compared to determine the quality of the approximations in terms of objective function, productions, energy stored, and dual variables (e.g., prices, storage value, and water value).

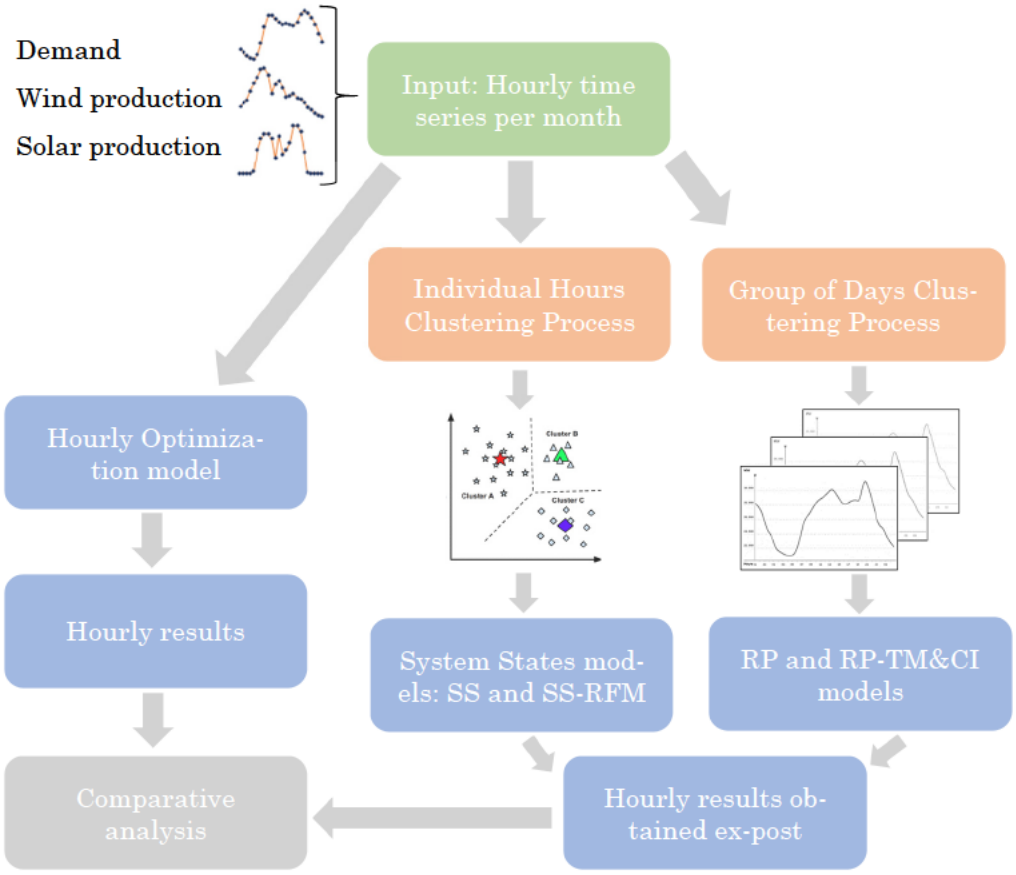


Figure 4-3. Analysis overview: comparison of LDC and LRP models

It is important to highlight that both load levels and representative periods are selected per month because the uncertainty in the hydro inflows is per month, see Figure 4-2. Therefore, the load levels and representative periods are different among months. This is guaranteed through a time division structure in both models. Figure 4-4 shows an example of the structure for the LDC model, where two months (m_1, m_2) have a subdivision in weekdays (w_1) and weekend (w_2), each one with three different load levels (l_1, l_2, l_3). Figure 4-5 shows an example of the structure for the LRP model, where the two months have their own representative periods (i.e., rp_1 and rp_2 for m_1, and rp_3 and rp_4 for m_2), each one with a set of chronological hours (k_1 to k_{24} in a 24-hour representative period example).

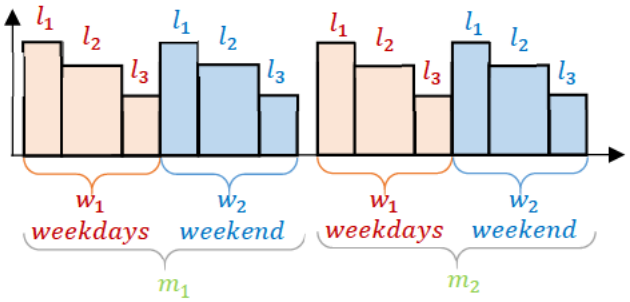


Figure 4-4. Structure of LDC time division

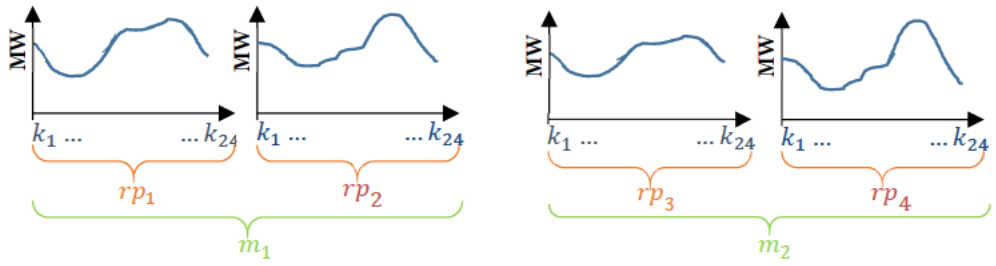


Figure 4-5. Structure of LRP time division

These time divisions facilitate the formulation of storage balance constraints in both models considering the scenario tree structure. Table 4-1 shows the storage balance constraints in the three optimization models. Although the detailed optimization models are shown in Appendix B.2 to B.4, the storage balance equations are shown to help with the analysis in the following sections.

Table 4-1. Storage Balance Constraints

Model	Storage Type	Formulation
HM	LTESS	$R_{p-1r}^{intra,\omega'} - R_{pr}^{intra,\omega} + i_{pr}^{\omega} - S_{pr}^{\omega} + \underbrace{\sum_{r' \in RUR_{r',r}} S_{pr'}^{\omega}}_{\text{Water spillage from upstream reservoirs}} + \underbrace{\sum_{h \in HUR_{h,r}} {P_{ph}^{\omega}}/{C_h}}_{\text{Turbined water from upstream hydro plants}} - \underbrace{\sum_{h \in RUH_{r,h}} {P_{ph}^{\omega}}/{C_h}}_{\text{Turbined water from hydro plants in reservoir } r} + \underbrace{\sum_{h \in HPR_{h,r}} {C_{ph}^{\omega}}/{C_h}}_{\text{Pumped water from hydro plants to reservoir } r} - \underbrace{\sum_{h \in RPH_{r,h}} {C_{ph}^{\omega}}/{C_h}}_{\text{Pumped water to other reservoirs}} = 0 \quad \forall \omega pr \;\; \omega' \epsilon\, a(\omega)$
HM	STESS	$SoC_{p-1b}^{intra,\omega'} - SoC_{pb}^{intra,\omega} - P_{pb}^{\omega} + C_{pb}^{\omega} = 0 \quad \forall \omega pb \;\; \omega' \epsilon\, a(\omega)$
LDC	LTESS	$R_{m-1r}^{inter,\omega'} - R_{mr}^{inter,\omega} + i_{mr}^{\omega} - S_{mr}^{\omega} + \underbrace{\sum_{r' \in RUR_{r',r}} S_{mr'}^{\omega}}_{\text{Water spillage from upstream reservoirs}} + \underbrace{\sum_{wl} \sum_{h \in HUR_{h,r}} {wg_{mwl} \cdot P_{mwlh}^{\omega}}/{C_h}}_{\text{Turbined water from upstream hydro plants}} - \underbrace{\sum_{wl} \sum_{h \in RUH_{r,h}} {wg_{mwl} \cdot P_{mwlh}^{\omega}}/{C_h}}_{\text{Turbined water from hydro plants in reservoir } r} + \underbrace{\sum_{wl} \sum_{h \in HPR_{h,r}} {wg_{mwl} \cdot C_{mwlh}^{\omega}}/{C_h}}_{\text{Pumped water from hydro plants to reservoir } r} - \underbrace{\sum_{wl} \sum_{h \in RPH_{r,h}} {wg_{mwl} \cdot C_{mwlh}^{\omega}}/{C_h}}_{\text{Pumped water to other reservoirs}} = 0 \quad \forall \omega mr \;\; \omega' \epsilon\, a(\omega)$
LDC	STESS	$SoC_{m-1b}^{inter,\omega'} - SoC_{mb}^{inter,\omega} - \sum_{wl} wg_{mwl} \cdot P_{mwlb}^{\omega} + \sum_{wl} wg_{mwl} \cdot C_{mwlb}^{\omega} = 0 \quad \forall \omega mb \;\; \omega' \epsilon\, a(\omega)$
LRP	Intra-period LTESS	$R_{rp,k-1,r}^{intra,\omega'} - R_{rp,k,r}^{intra,\omega} + i_{rp,k,r}^{\omega} - S_{rp,k,r}^{\omega} + \underbrace{\sum_{r' \in RUR_{r',r}} S_{rp,k,r'}^{\omega}}_{\text{Water spillage from upstream reservoirs}} + \underbrace{\sum_{h \in HUR_{h,r}} {P_{rp,k,h}^{\omega}}/{C_h}}_{\text{Turbined water from upstream hydro plants}} - \underbrace{\sum_{h \in RUH_{r,h}} {P_{rp,k,h}^{\omega}}/{C_h}}_{\text{Turbined water from hydro plants in reservoir } r} +$

Model	Storage Type	Formulation
		$\underbrace{\sum_{h \in HPR_{h,r}} C^{\omega}_{rp,k,h}/c_h}_{\text{Pumped water from hydro plants to reservoir } r} - \underbrace{\sum_{h \in RPH_{r,h}} C^{\omega}_{rp,k,h}/c_h}_{\text{Pumped water to other reservoirs}} = 0 \quad \forall \omega, rp, k, r \quad \omega' \epsilon\, a(\omega)$
LRP	Inter-period LTESS	$R^{inter,\omega'}_{m-1,r} - R^{inter,\omega}_{mr} + \sum_{(rp,k)\in\{CI_{p,rp,k} \cap MP_{m,p}\}} \left[R^{intra,\omega}_{rp,k,r} - R^{intra,\omega'}_{rp,k-1,r}\right] = 0 \quad \forall \omega mr \quad \omega' \epsilon\, a(\omega)$
LRP	Intra-period STESS	$SoC^{intra,\omega'}_{rp,k-1,b} - SoC^{intra,\omega}_{rp,k,b} - P^{\omega}_{rp,k,b} + C^{\omega}_{rp,k,b} = 0 \quad \forall \omega, rp, k, b \quad \omega' \epsilon\, a(\omega)$
LRP	Inter-period STESS	$SoC^{inter,\omega'}_{m-1,b} - SoC^{inter,\omega}_{mb} + \sum_{(rp,k)\in\{CI_{p,rp,k} \cap MP_{m,p}\}} \left[SoC^{intra,\omega}_{rp,k,r} - SoC^{intra,\omega'}_{rp,k-1,r}\right] = 0 \quad \forall \omega mb \quad \omega' \epsilon\, a(\omega)$

Where indices r, h, b, p, and ω represent the elements in sets for reservoirs, hydro units, STESS (e.g., batteries), periods (e.g., hours), and scenarios, respectively. Lowercase terms are used for parameters while uppercase terms denote variables. For instance, i^{ω}_{pr} is the hydro-inflows parameter for each reservoir r and period p at scenario ω. Moreover, $R^{intra,\omega}_{pr}$ and $SoC^{intra,\omega}_{pb}$ are variables for the intra-period reservoir level and state-of-charge for each reservoir r and STESS b in period p at scenario ω. The complete list of parameters and variables is detailed in Appendix B.1. In the *HM model*, the storage balance constraints are imposed for each period p. Therefore, reservoir level and SoC are determined for each hour in the time horizon. These results are used as benchmark to test the LDC and LRP model, see Figure 4-3.

Constraints for LTESS and STESS are stated for $\omega' \epsilon\, a(\omega)$, which allows to relate the different scenarios through the scenario tree. For instance, Figure 4-2 shows a scenario tree with three scenarios: wet season, average inflows, dry season. In this example, the ancestor $a(\omega)$ of scenario 3 in November is scenario 1 in October. Therefore, the set $a(\omega)$ is relating a scenario with the corresponding predecessor scenario in the tree.

In the *LDC model*, both storage balance equations (i.e., LTESS and STESS) include the load-level duration (wg_{mwl}) to consider the number of hours that are represented for each load level. In other words, the multiplication by wg_{mwl} guarantees that all the charged/discharged energy is considered within the month m. These equations are for the inter-period variables. Intra-period variables are not available in this model due to the lack of chronology within the month m.

In the *LRP model*, the storage balance constraints are defined for inter- and intra-periods. These equations create the continuity in storage across the entire time horizon that allows for the modeling of short-term and long-term storage simultaneously. Intra-period constraints ensure the storage balance within the RP, while inter-period constraints guarantee the storage balance by checking at regular intervals (e.g., aggregation of hours such as months m) that all the energy charged and discharged since the previous month plus the total energy at the last checkpoint are within bounds. This is possible because the cluster index, $CI_{p,rp,k}$, and the relationship between periods and months, $MP_{m,p}$, are known as a result of the clustering procedure to determine the RPs,

see Section 3.2.4. The intersection of both sets $\{CI_{p,rp,k} \cap MP_{m,p}\}$ indicates which RPs belong to the month and, therefore, must be considered in the inter-period balance.

Finally, notice that constraints for LTESS and STESS are equivalent if, for example, a hydro reservoir has a pump unit which is not in a hydro basin and it has no hydro inflows. However, we keep both constraints in order to facilitate the distinction between both types of storage technologies. In real hydro power plants, there is a nonlinear dependence between the reservoir head and the reservoir volume [112]. Nevertheless, and for the sake of simplicity in storage balance constraints for LTESS, we assume a linear function of the turbine outflow. Although nonlinear dependence could be considered at the expense of more complex optimization models such as in [112].

4.4 Analysis of Energy Storage Opportunity Cost

In Appendix B.2 to B.4, we formulated three models for the hydrothermal dispatch as Mixed Integer Programming (MIP) problems. As it is mentioned in [113], the value of dual variables in a MIP are not well-defined Hence, we apply the common practice to approximate the dual variables of interest by fixing the integer variables (e.g., commitment decisions) obtained in the MIP solution and then solving the model again as a Linear Programming (LP) problem [112]. Under this assumption, the opportunity cost of ESS can be obtained from the dual variable of the storage balance equations of each model, while the opportunity cost of hydro reservoirs is normally called water value [113]. However, the name *water value* cannot be applied to BESS since there are no hydro inflows for this type of technology. Instead, we use *storage value* to describe the opportunity cost of short-term storage (i.e., BESS).

Hourly Model (HM): The opportunity costs for each type of storage are obtained from the dual variables of LTESS and STESS constraints in Table 4-1. Therefore, water value $\left(\mu_{pr}^{intra,\omega}\right)$ is obtained from HM-LTESS constraint and storage value $\left(\mu_{pb}^{intra,\omega}\right)$ from HM-STESS constraint. These opportunity costs are for each hour in the time horizon.

Load Duration Curve Model (LDC): The water value $\left(\mu_{mr}^{inter,\omega}\right)$ is obtained from LDC-LTESS in Table 4-1 and storage value $\left(\mu_{mb}^{inter,\omega}\right)$ from LDC-STESS in Table 4-1. Since there is no chronology between load levels, the opportunity cost for each type of storage is obtained only for an aggregation of hours (e.g., months).

Linked Representative Periods Model (LRP): This model has two balance equations for each storage technology. One for the storage balance inside the representative period (intra-period) and another for the storage balance through the aggregation of hours in the time horizon (inter-period). Each balance equation has its dual variable; however, the combination of both dual variables is necessary to determine the equivalent hourly dual variable from the HM model. Equation (4-1) defines the hourly storage/water value for short- and long-term storage using the LRP model. $\mu_{rp,k,s}^{intra,\omega}$ is obtained from the dual variables of intra-period constraints in Table 4-1. In the same way, $\mu_{m,s}^{inter,\omega}$ is obtained from the dual variables of inter-period constraints in Table

4-1. Equation (4-1) shows the opportunity cost of energy storage as a linear combination of short- (intra-period balance) and long-term decisions (inter-period balance). Therefore, the LRP model distinguishes the impact of short-term decisions within the total opportunity cost, which is not possible in the HM model.

$$\mu_{ps}^{intra,\omega} = \sum_{(rp,k)\in CI_{p,rp,k}} \sum_{m\in MP_{m,p}} \frac{1}{p_m^{\omega}} \cdot \left(\frac{\mu_{rp,k,s}^{intra,\omega}}{wg_{rp}} + \mu_{m,s}^{inter,\omega} \right) \quad \forall \omega ps \tag{4-1}$$

4.5 Case Study and Results

As a case study, we chose a stylized Spanish power system in target year 2030. The Spanish case is relevant because it has hydro reservoirs (i.e., seasonal storage) and, according to ENTSO-E [100], the next ten years will likely bring investment in Battery Energy Storage System (BESS), i.e., short-term energy storage. The wind and solar profiles were taken from [101], [102] while hourly demand data and annual production per technology were taken from the scenario 1 in the ENTSO-E *Ten Year Network Development Plan 2016* [100].

The water basin is represented by three reservoirs. Reservoirs 1 and 2 are upstream of reservoir 3, and, therefore, reservoir 3 receives, besides its hydro inflows, the hydro production and water spillage from reservoirs 1 and 2, such as in Figure 4-1. The scenario tree is a simplified structure of three scenarios in order to consider monthly hydro inflows in dry, average, and wet seasons, see Figure 4-2. The probabilities for each scenario are 30%, 50%, and 20%, respectively. The first-stage decision is taken for October[12] and second-stage decisions are taken from November to September. For the sake of simplicity, the stability of the solution for different scenario trees is not verified. This will be addressed in future research to determine the impact of different scenario trees in the results.

Load levels and representative periods are obtained via the k-means clustering procedure for each month. The clusters were chosen by normalizing time series for the hourly demand, wind availability, and solar availability, see Figure 4-3. We have defined 12 load levels (6 for weekdays and 6 for the weekend) per month for the LL model. For the LRP model, we defined some sensitivities for the selection of the representative periods: 1 representative period with 24h per month (1RPx24h), 1 rp with 48h per month (1RPx48h), 1 rp with 96h per month (1RPx96h), 2 rp with 24h per month (2RPx24h), and 4 rp with 24h per month (4RPx24h). These sensitivities are performed in order to identify if it is better to have only one rp per month sharing information or to have more rp per month sharing information among them and between months. Based on previous results in [95], we expect that more rp per month sharing information is better than one rp per month.

[12] October is the beginning of the hydrological year in Spain.

Finally, we consider a BESS with a power rating of 200 MW, energy capacity of 4 hours, and round-trip efficiency of 90%. The BESS is installed to deal with hourly variation of variable renewable energy sources.

4.5.1 Objective Function and CPU Time

Table 4-2 shows the results using as a reference the results obtained for the HM. The objective function error is calculated using the value of the objective function of the hourly model as the theoretical value, while the CPU time is shown as a fraction of the time taken by the hourly model to solve the problem. All models were solved until optimality, i.e., until either their optimal point or the integrality gap equaled zero.

The analysis of the results shows two main situations. First, the LRP 4RPx24h as the best performance in terms of the objective function. The objective function error is lower than 1% compared to the HM model, and it only takes one tenth of the time to solve. In addition, all the LRP results for the different sensitivities have objective function errors lower than 4%. Second, although the LDC is one of the fastest models to solve the problem, its objective function error exceeds 10%[13]. Therefore, the LRP improves the results of the hydrothermal dispatch problem without hampering the computational efficiency.

The results obtained for the sensitivities of the LRP model confirm that more *rp*s per month sharing information is better than one longer *rp* per month. For instance, the 1RPx48h and 2RPx24h take the same number of hours per month and produce similar CPU time performance. However, the objective function error in 2RPx24h is half of that obtained with 1RPx48h. Therefore, and for the sake of simplicity, we show only the LRP 4RPx24h model results in the following sections.

Table 4-2. Objective Function Error and CPU Time

	LRP 4RPx24h	**LRP 2RPx24h**	**LRP 1RPx96h**	**LRP 1RPx48h**	**LRP 1RPx24h**	**LDC**
OF Error [%]	0.1	1.7	3.4	3.6	3.6	11.7
CPU Time [p.u.]	0.10	0.02	0.05	0.02	0.01	0.01

4.5.2 First- and Second-Stage Production Results

Table 4-3 shows the errors in production per technology for both LDC and LRP model. Negative values indicate an underestimation in comparison to the HM result, while positive values indicate an overestimation. The black color is used to highlight absolute values lower or equal to 5%, light orange color for absolute values greater than 5% and lower or equal to 10%, and the dark red color for absolute values higher than 10%. Technologies such as coal and fuel oil are not shown because their total production is negligible. Additionally, technologies such as wind, solar, and run-of-river are also not shown, in this case because the total production error is lower than 1% in both models.

[13] Considering twice the number of LL, the LDC objective function error reduces to an error of 6%, but the CPU time increases six-fold.

The results are classified into two groups: first and second stage. The LRP model has better results than LDC model in both groups. In fact, BESS production error in the LRP model is almost ten times lower than the result with the LDC model for the first stage, and almost twenty times smaller for the second stage. The Open Cycle Gas Turbine (OCGT) is more difficult to estimate in both models because it is the peak technology, and yet the LRP improves the approximation made by the LDC.

Table 4-3. Total Production Error per Technology [%]

Tech	First Stage		Second Stage	
	LDC	LRP	LDC	LRP
Nuclear	4.6	1.7	Sc1 2.9 Sc2 2.8 Sc3 3.3	Sc1 -0.3 Sc2 -0.4 Sc3 -0.7
CCGT	-35.8	-0.5	Sc1 -8.6 Sc2 -5.7 Sc3 -6.8	Sc1 -3.5 Sc2 -4.4 Sc3 -2.6
OCGT	-30.7	-22.7	Sc1 -68.4 Sc2 -44.3 Sc3 -68.8	Sc1 -49.0 Sc2 -0.5 Sc3 -16.9
Hydro	-5.5	-4.3	Sc1 -0.4 Sc2 0.6 Sc3 -0.3	Sc1 -0.3 Sc2 -0.5 Sc3 -0.3
BESS	110.9	9.1	Sc1 133.9 Sc2 115.1 Sc3 123.7	Sc1 -5.4 Sc2 -5.1 Sc3 4.6

The LDC model underestimates Combined Cycle Gas Turbine (CCGT) production (marginal technology most of the time) due to the loss of chronology among the load levels that overestimates BESS production, see Table 4-3. This leads to an underestimation of BESS storage values, see Table 4-3. On the other hand, the CCGT production error in the LRP model is lower than 5% as well as the BESS storage value in Table 4-4 thanks to the more accurate representation of chronological constraints. These results show the interdependence between the marginal technologies and the BESS storage value.

4.5.3 Hydro Reservoir and SoC Results

In this section, we analyze the storage level for both technologies, i.e., hydro (seasonal storage) and BESS (short-term storage). First, the storage level for hydro reservoir is approximated with more accuracy in the LRP model compared to the LDC model. For instance, Figure 4-6 (top) shows the storage level for reservoir 2 in scenario 1 for both models in comparison to the HM model. The storage-level-average error over all the scenarios and reservoirs is 4.5% for the LRP model and 9.0% for the LDC model. Therefore, the LRP model is twice as accurate as the LDC model for the reservoir levels.

Second, the hourly BESS SoC can be obtained for the LRP model, which is not possible with the LDC model. Figure 4-6 (middle) shows the hourly evolution of the SoC in a particular week for the HM and the LRP model. We can

observe the daily cycles of the BESS and how the LRP model results mimic the HM solution. We also want to compare the BESS results obtained from the LDC model. Therefore, we have estimated the total number of cycles[14] obtained from each model in the target year. Figure 4-6 (bottom) shows the total number of cycles per scenario for each model. The total number of cycles determined from the LDC results doubles the number obtained with the HM model, which was expected due to the overestimation in the BESS production shown in Table 4-3. By contrast, the average error in the number of cycles for the LRP model is 5%. This result is important because the number of cycles is key to determine replacement or maintenance in BESS.

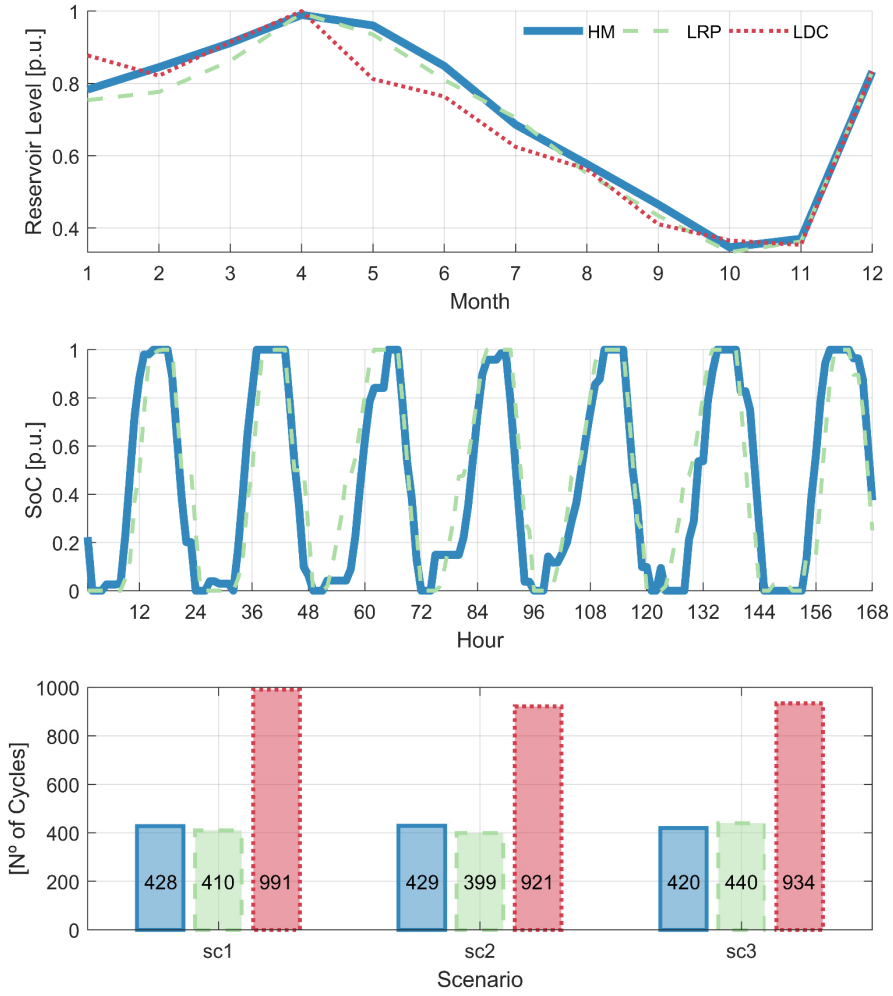


Figure 4-6. Reservoir Level, BESS SoC, and BESS total number of cycles

4.5.4 Marginal and Opportunity Costs

Table 4-4 shows the errors with Stochastic Marginal Cost (SMC) and Opportunity Cost (OP). The HM results are chosen as a reference for the error calculation. The SMC is calculated as the weighted dual variable from the balance equation in each model. The OP is calculated as the weighted dual variable from the inter-period storage balance equation in each model in Table 4-1. We use the same color notation as in Table 4-3. On the one hand, the LRP

[14] The cycles are estimated for all models using the total charge/discharge energy over the year and dividing it by the BESS' maximum energy capacity.

model mostly leads to errors lower than 5% and is the most accurate model in almost all results. On the other hand, the LDC model yields in most of the cases errors higher than 10% and, as expected from the results in previous sections, exhibits the worst performance in the opportunity cost of the BESS throughout the time horizon.

The errors in reservoir 3 are higher than those in reservoir 1 and 2. This is a reasonable result, considering that reservoir 3 is downstream of reservoir 1 and 2. Therefore, errors in reservoirs 1 and 2 propagate to reservoir 3 and complicate the estimation.

Table 4-4. Stochastic Marginal Cost and Opportunity Cost – Error [%]

Month	SMC		OP Res 1		OP Res 2		OP Res 3		OP BESS	
	LDC	LRP	LDC	LRP	LDC	LRP	LDC	LRP	LDC	LRP
Oct	-13.7	1.3	-16.3	-0.6	-6.6	-0.6	-23.9	2.8	-6.7	-2.0
Nov	-4.7	-2.0	-16.3	-0.6	-6.6	-0.6	-23.9	2.8	-4.7	-0.4
Dec	-28.5	0.3	-16.3	-0.6	-6.6	-0.6	-23.9	2.8	-16.3	-3.6
Jan	-10.8	-1.3	-17.7	-2.4	-8.2	-2.4	-22.6	-1.2	-32.2	-0.9
Feb	-8.7	0.7	-12.5	-0.4	-8.2	-0.4	-22.6	-1.0	-32.4	-3.1
Mar	-13.5	6.6	-12.5	-0.4	-8.2	-0.4	-21.4	-1.0	-27.4	5.6
Apr	-9.5	0.6	-12.5	-0.4	-8.2	-0.4	-21.4	-1.0	-29.0	-3.2
May	-14.0	0.2	-12.5	-0.4	-8.2	-0.4	-21.4	-1.0	-26.7	0.0
Jun	-9.4	1.4	-12.5	-0.4	-8.2	-0.4	-21.4	-1.0	-26.3	4.3
Jul	-23.1	2.0	-12.5	-0.4	-8.2	-0.4	-21.4	-1.0	-37.9	4.3
Aug	-13.8	4.7	-11.9	1.2	-6.9	-0.6	-22.3	-1.4	-30.2	1.9
Sep	-15.0	1.5	15.4	0.4	-6.9	-0.6	6.0	0.3	-28.8	-0.9

In Section 4.4, we have defined the equation (4-1), which allows us to determine the intra-period or hourly opportunity cost in the LRP model. This represents the main advantage over the LDC. Figure 4-7 shows the opportunity cost (or storage value) of BESS in the HM, LRP, and LDC models for a particular week. The opportunity cost obtained from the LRP model mimics the trend followed by the results in the HM model. In fact, almost 75% of the time, the difference between the results of both models is lower than 10%. Figure 4-7 shows one value for the LDC model throughout the week because it only determines one opportunity cost value per month, see Table 4-1.

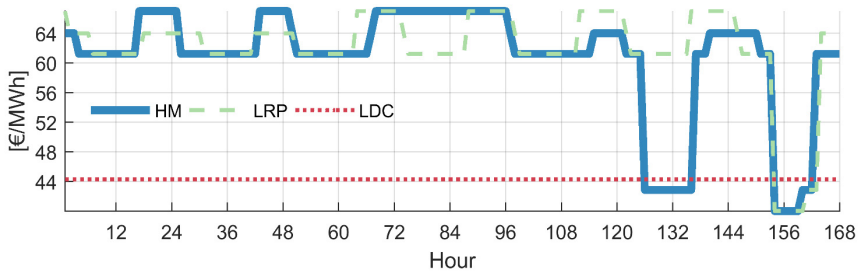


Figure 4-7. Opportunity Cost or Storage Value of BESS [€/MWh]

4.6 Discussion

The coordination of short- and medium-term hydrothermal models has been traditionally performed by using two models and sharing information between them [114]–[116]. One medium-term model runs first using a reduced chronology, such as the LDC model described in Section 4.3, in order to obtain the end volume or the water value from each reservoir. Under the assumption that a Stochastic Dual Dynamic Programming (SDDP) [114], [117] approach has been used to solve the medium-term model, a piecewise-linear Future Cost Function (FCF) can be utilized to meet end-point conditions from the medium-term model in the short-term model [118]–[120]. This information is used as input data in a short-term model to find the daily levels and hourly economic signals.

The main drawback of this process is that the medium-term model, i.e., LDC model, does not consider short-term chronological information. In fact, as shown in Section 4.5.1, the LDC model has the worst performance. All the time resolutions tested for the LRP model have shown a better performance than the LDC model. This means that LRP succeeds in the internalization of short-term chronological information in the medium-term hydrothermal problem, which enables inclusion of the operation of BESS without solving the HM model for the entire medium-term horizon. In other words, the LRP model co-optimizes both medium- (or long-) and short-term decisions.

The LRP model is also compatible with decomposition techniques, such as Benders' decomposition or SDDP, in order to consider a large number of scenarios in the scenario tree. Therefore, we could obtain the FCF internalizing the hourly dynamics of short-term storage, which is not possible with the current LDC model approach.

4.7 Main Takeaways in this Chapter

Through this chapter we have proposed a new formulation to include short-term ESS operational decisions in a stochastic hydrothermal dispatch model, which is based on the proposed enhanced version of the representative periods approach that allows an analysis of both short- and long-term storage at the same time. The main analysis and results can be summarized in the following takeaways:

- *We have validated the initial hypothesis that short-term energy storage (e.g., BESS) decisions on energy production impact the opportunity cost (or water value) of seasonal storage.*
- *For the analyzed case study, the classical LDC approach underestimates the water value to be between 6% and 24% for seasonal hydro reservoirs.*
- *Operational results (e.g., productions, number of cycles for short-term storage, and storage levels) have a better estimation in the LRP model than in the LDC model.*
- *Hourly opportunity costs internalize long-term signals due to seasonality in the power system. In other words, the water value in seasonal storage includes the impact of short-term operational decisions.*

- *This is important to help market participants or planning authorities in their decision-making process (bids or investment decisions) by determining correct economic signals (i.e., short-term prices and long-term expected values) with the co-optimization approach in the LRP model and therefore avoiding sub-optimal solutions from iterative processes (e.g., fixing the hydro reservoirs levels obtained from a medium-term model in a short-term operational model).*

5 ESS INVESTMENT MODELING

Contents

This chapter studies optimization models for ESS investment in medium- and long-term planning. The results are divided in two main sections. The first section investigates the effects of transmission losses, constraints and increased renewable energy penetration on planning ESS allocation and investment. By modifying a DC Optimal Power Flow model using a linearized approximation for ohmic losses we were able to understand which network characteristic or inhibitor drives the most change in expanding utility scale storage. The results show that network congestion is a more influential network inhibitor than ohmic losses representation. Ohmic losses representation only has substantial effects on an uncongested network and produce marginal changes in ESS allocation when in congested ones.

The second section compares the system-states type and representative-periods type methods when short-term energy storage and seasonal storage are included in the optimization models for ESS investment decisions. As in Section 3.2, the proposed enhanced version of the representative periods method succeeds in a more accurate representation of both short- and long-term storage, leading to almost 10 times lower error in ESS investment results in comparison to the other models analyzed. Finally, we analyze the close relationship between renewable curtailment and ESS investment in both type of models. The results show that models in which the renewable curtailment is underestimated, the ESS investment is underestimated as well.

The analysis and results in this section were published in Article III [121] *and Article II* [95].

5.1 Analysis of Transmission Network in ESS Investment

Designing a storage network on a grid-scale requires examining capacity requirements, resource allocation, and grid-specific properties [50]. A common method used to account for this is the optimal power flow (OPF) framework. Studies [122], [123] have explored adding charge and discharge dynamics to AC OPF functions in order to explore the effects of energy trading, demand reduction and power regulation. Others e.g., [124], [125] have looked at expanding this type of modeling to include multiperiod storage location optimization. Studies such as [126] have also used AC OPF models to study the economic benefits realized using energy storage for emissions reductions and/or congestion relief.

In general, OPF Problems are nonconvex and NP-hard making them difficult to solve. Therefore, DC OPF linear approximations [127] as opposed to the AC OPF models mentioned above, have been used extensively to explore optimal storage siting problems for both customary and renewable energy grids. Much attention has been brought to using DC OPF Functions to fluctuation challenges associated with high renewable penetration [128]. These models have also been used to understand how storage can be used as a risk mitigating measure regarding the uncertainty of renewable generation [129].

The choice of model for this analysis was based on the following: General transportation models only consider Kirchhoff´s First Law of energy/power conservation which is not a viable means of planning technology expansion. Contrarily DC power flow models (DCPF) consider both Kirchhoff's First Law and a linearized version of his second law to account for voltage balance. These models allow users to approximately represent losses but not stability considerations in transmission. This makes them suitable for medium and long-term planning for investment and allocation of new elements in the power grid (i.e. generators, transmission lines, transformers, etc.) [130]. As this study is not concerned with exploring voltage and stability problems associated with grid operation, building upon the DC OPF-based model in [50] was sufficient. This allows for increased computational tractability that would be lost had this model been expanded to an AC power flow representation. [50] expanded on the DC OPF-based storage allocation formulas to include a portfolio of storage technologies operating on different time-scales into the OPF-based siting and dispatch problems. This was used to optimally allocate and invest in storage in both congested and uncongested networks.

Congestion can strongly impact the sizing and siting of storage facilities by preventing proper power distribution, leading to outages or increased system stress. Losses in power flow throughout a network can similarly affect grid reliability by preventing the right amount of power from reaching demand centers. Studies such as [131] have explored how energy storage technologies can improve overall grid efficiencies by relieving congestion, stabilizing power and minimizing transmission and BESS round-trip losses. While this is crucial in effectively integrating distributed storage solutions, it neglects to give a system planner an optimal network-level strategy for designing new storage projects. Ignoring line losses may initially show lower total system costs, however, it can lead to expensive investment adjustments down the road due to disregarding transmission losses [132]. This work expands on [50] to account

for both network congestion and transmission losses. A linearized approximation of ohmic losses was adapted to do this [133]. By investigating physical constraints in a transmission network, the impact that storage expansion will have on grids operating with non-dispatchable power can be better analyzed.

5.1.1 Optimization Models

In order to analyze the impact of transmission losses and congestion in ESS *allocation* and *investment*, two optimization model are defined. Appendix C details both models, one for the allocation analysis and another for the investment problem. Moreover, the models are an extension of the OPF problem with chronological constraints (e.g., ramping constraints, energy storage balance). It includes multiple options for storage technologies, creating a larger problem size. To increase the tractability of the problem a linearized DC OPF approximation was adopted. The *allocation* model fixes the total storage capacity available for the mix of technologies and optimizes the sitting of these resources throughout the network while considering linearized ohmic losses. The *investment* model further extends the allocation model to include investment costs for installing storage capacity. Section 5.1.2 breaks down how the linearized approximation and subsequently its related terms and constraints are formulated.

Both models in Appendix C aim to analyze the allocation and investment decisions for STESS only. Section 5.2 analyzes the ESS investment problem while considering both STESS and LTESS at the same time.

5.1.2 Transmission Losses Modeling

The key attribute of this analysis is the consideration of losses throughout the system. To maintain the optimization model's tractability, a linearized piecewise ohmic loss approximation is used [133]. Doing this requires starting with cosine and quadratic estimations of transmission losses. First, a cosine curve was used to represent the active power dissipated in a circuit. This provides a good approximation of transmission losses. While the function itself is convex, the problem remains non-convex and non-linear preventing a globally optimal solution from being found. To address this, we first use a Taylor Series expansion of the cosine function to obtain a quadratic approximation of transmission losses.

The resulting losses can be estimated to the first two elements of the expansion since the subsequent terms yield negligible values. This is a relatively standard method of simplifying the optimization problem. However, it results in the same problem as before—the model remains non-convex and non-linear. To correct for this, the quadratic approximation can be estimated using a linearized piecewise function [133]. Figure 5-1 shows the linear piecewise approximation relative to the quadratic representation, where index k denotes the linear segments in the piecewise function, $m_{t,k}$ is the slope for each linear segment k at time t, $\Delta_{nm}(t)$ is voltage angle difference between nodes n and m at time t, and $\Delta'_{nm}(t)$ is the approximated quadratic function for transmission losses.

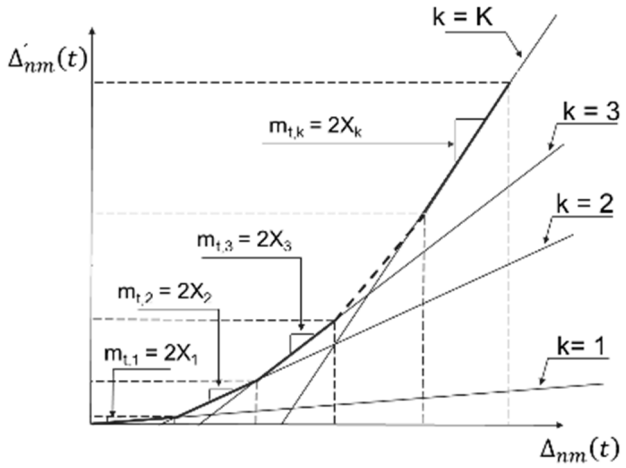


Figure 5-1. Piecewise Linear Approximation of Ohmic Losses

In addition, two new parameters are introduced in order to perform the linear approximation: X_k Original x-axis values used, and I_k y-intercept values for each segment. These parameters allow us to calculate the equations of each segment used to approximate the quadratic function with linear curves at specific points via (5-1), where subset Θ_n guarantees that only connected nodes through a transmission line are considered. Using them we find that the values of $\Delta'_{nm}(t)$ are approximately equal to the square of the voltage angle differences used in the general quadratic approximation. Therefore, transmission losses on a line can then be represented as $0.5\left(Gt_{nm}\Delta'_{nm}(t)\right)$, where Gt_{nm} is the circuit conductance. Note that this product is halved as to prevent counting both the positive and negative sides of the approximated curve.

$$\Delta'_{nm}(t) \geq \left(2X_k\Delta_{nm}(t)\right) + I_k \qquad \forall k, m \in \Theta_n, t \quad (5\text{-}1)$$

With this approximation, the optimization problem becomes both linear and convex. It can therefore be solved to obtain a globally optimal solution. Generally, it is possible that the optimal solution could choose a fictitious amount of system losses (e.g., a point above the approximation curve). This is most likely to occur in a unit commitment problem, where the model would find it more economically favorable to artificially increase transmission losses to prevent a thermal generator from incurring relatively large shut-down costs. But it is not physically possible to make a transmission line more resistive than its material composition allows for. To correct this, binary variables can be used to restrict power loss values to the piecewise function, thereby creating a MIP, and consequently an increasingly complex problem [134]. Because this analysis did not deal with unit-commitment decisions, this additional complication was unnecessary.

5.1.3 ESS Performance Metrics

To best comprehend model output and draw conclusions on allocation and investment decisions, three additional metrics were calculated using the aforementioned parameters. These metrics are defined as follows:

$$OM_{jn} = 1 - \frac{max_t\{s_{jn}(t)\}}{k_{jn}} \qquad \forall j, n \qquad (5\text{-}2)$$

$$CM_{jn} = \frac{\sum_{t \in T}\left[r_{jn}^{c}(t) \cdot \Delta t\right]}{k_{jn}} \qquad \forall j,n \ \ if \ k_{jn} > 0 \qquad (5\text{-}3)$$

$$OSL_{j} = \frac{\sum_{n,t} s_{jn}(t)}{\sum_{n,t} s\_BaseCase_{jn}(t)} \qquad \forall j \qquad (5\text{-}4)$$

Where index j denotes the set of energy storage technologies, $s_{jn}(t)$ represents the storage level, k_{jn} represents the installed storage capacity, $r_{jn}^{c}(t)$ correspond to the charging rates of each storage technology, and Δt is the duration of time step.

The Overall Capacity Metric (OM) in (5-2) compares the maximum storage level in MWh attained over the time horizon to the actual amount of capacity of that technology installed at each node. The Cycling Metric (CM) in (5-3) keeps track of how many full charging cycles a technology goes through over the total time horizon at each node. Lastly, the Overall Storage Level Metric (OSL) in (5-4) provides an idea of how much energy each technology stores throughout a day for each scenario in comparison to the base case of an unconstrained network.

5.1.4 Case Study Description

The numerical examples are based on the 14-bus IEEE benchmark system [135]. Transmission constraints are adopted into this model via TC_{nm}^{max} parameter. Unless otherwise stated the transmission capacity between any two nodes is 400 MW. Note that all data inputs (excluding those relevant to the losses formulation) are adopted from [50].

This includes information regarding the charge/discharge rates [8] and total storage capacity and charge/discharge efficiencies [99]. Relevant information regarding the available storage portfolio is presented in Table 5-1.

Table 5-1. Storage Technology Parameters

Storage Technology	Dis/Charge Efficiency η^{c}, η^{d} [$p.u.$]	Investment Cost C_j^i [$/MW per day]	Discharge Duration DD_j [h]
PSH	0.87	250	12
CAES	0.78	24	24
BESS	0.94	800	4
FES	0.96	550	1/12

Four different storage technologies were considered: pumped-storage hydro (PSH), compressed air energy storage (CAES), battery energy storage system (BESS), and flywheel energy storage (FES). Studies were run over a 24-hr time horizon with time steps Δt = 5-min throughout the day allowing for both long- and short-time scale observations. Both conventional thermal and wind generation occur at buses 1, 2, 3, 6 and 8. Wind production is treated as a parameter. As such, the model is set up to force all wind generated power to be accepted. The information is adopted from the 2006 NREL Western Wind Resources Dataset [136]. Wind data was available in 10-min intervals. A set of interpolated 5-min intervals was used to run these cases.

In Sections 5.1.7 and 5.1.8, the wind generation was scaled up and down by fixed multipliers to represent different generation scenarios. The optimization models are developed in GAMS 25.1.1 and solved using the commercial

solver GUROBI 8.0. The solver defaults settings were used for all the experiments, which were run on an Intel-i7 CPU@3.4-GHz computer with 16GB of RAM memory and four cores.

5.1.5 Results for ESS Allocation in an Unconstrained Network

Here we analyze the allocation of storage in an unconstrained network with losses. We assume in this scenario that the maximum amount of capacity for each storage technology will always be allocated throughout the network.

Excluding losses, we find that the locational marginal pricing (LMP) at every node is the same. No technology is favored, and the capacity of each technology is distributed evenly system wide. However, no CAES storage was chosen at all. The absence of this technology is likely the result of its low round-trip efficiency. The introduction of losses into the model led to changes in the spatial distribution of storage capacity and the temporal usage of each technology. FES and PSH technologies exhibited the greatest changes. We can observe the changes in nodes 3 and 9 to best understand this.

Table 5-2. Storage Allocation, Additional Storage Metrics

Case Study	Tech	CM [p.u.]		OM [%]		OSL [p.u.]
		n3	n9	n3	n9	-
Congestion = NO Losses = NO	PSH	0.003	0.003	100.0	100.0	1.000
	CAES	0.003	0.003	100.0	100.0	0.000
	LION	1.734	1.734	92.9	92.9	1.000
	FES	5.305	5.305	92.9	92.9	1.000
Congestion = NO Losses = YES	PSH	0.004	0.001	100.0	100.0	1.242
	CAES	0.000	0.000	100.0	100.0	0.000
	LION	3.119	1.044	88.6	98.0	0.999
	FES	18.091	10.661	70.9	86.7	0.999
Congestion = YES Losses = NO	PSH	0.127	0.000	99.2	100.0	15.525
	CAES	0.000	0.000	100.0	100.0	0.000
	LION	1.174	2.110	9.4	100.0	1.071
	FES	6.640	0.000	24.5	100.0	1.111
Congestion = YES Losses = YES	PSH	0.090	0.000	99.2	100.0	15.797
	CAES	0.000	0.000	100.0	100.0	0.000
	LION	1.040	0.000	10.8	100.0	1.073
	FES	6.178	0.000	25.8	100.0	1.095

As expected, the CM in Table 5-2 indicates that FES goes through the most full-cycles of any technology in the portfolio. This is likely attributed to its short ramp time which allows it to fully charge and discharge within a single timestep. When losses are introduced, the number of cycles more than triples at node 3 and doubles at node 9 even though the overall capacity remains unchanged. FES is the *"fastest"* technology allowing it to stabilize short term fluctuations in load caused by wind generation. This characteristic becomes increasingly advantageous when there is less flexibility in power transmission to ease load volatility. PSH is a slow-moving technology but its large allowable capacity permits it to employ a day/night arbitrage pattern of operation. The OSL in Table 5-2 illustrates this trend with an increase of 24% in PSH storage usage midday, allowing a system operator to dissipate stored energy at the evening peaks. Looking at Figure 5-2 we see this pattern of

storage deployment illustrated by the step down in network-wide PSH storage level starting around time-step 175. The storage level remained at a plateau midday when demand was lower in order to be discharged when it was needed most.

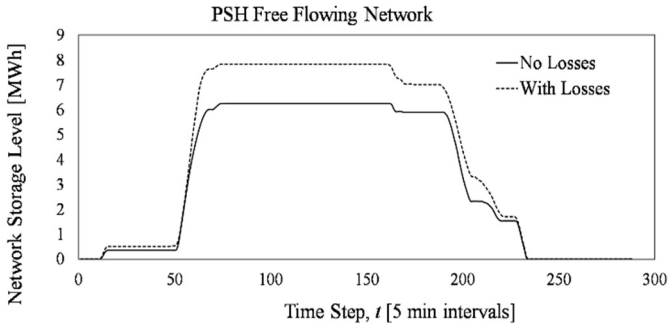


Figure 5-2. PSH Storage level in a free-flowing network

5.1.6 Results for ESS Allocation in a Constrained Network

In this section we explore how network congestion impacts storage allocation. Originally values representing a very congested network were used to simulate the case of *"full"* congestion. Congestion was introduced into the network as follows: $TC_{12}^{MAX} = 80MW, TC_{15}^{MAX} = 40MW, TC_{23}^{MAX} = TC_{42}^{MAX} = 30MW$. The remaining lines had a capacity of 400 MW.

Storage in an uncongested network exhibited significant changes in location and usage when losses were introduced, however this was not the case for the constrained case. In this scenario, a system operator is limited as to where they can route power to meet demand. This makes the system unable to deal with rapidly changing power flows such as those discharging from flywheels. The CM and OM metrics at bus 3 for FES in Table 5-2 indicate this. The CM did not change significantly, and the OM dropped to 24%, indicating that FES storage was not used at full capacity in the constrained network. Considering losses did not change these values significantly. For PSH storage network congestion increased the OSL fifteen-fold system-wide from the base case. Again, considering losses yielded no significant changes.

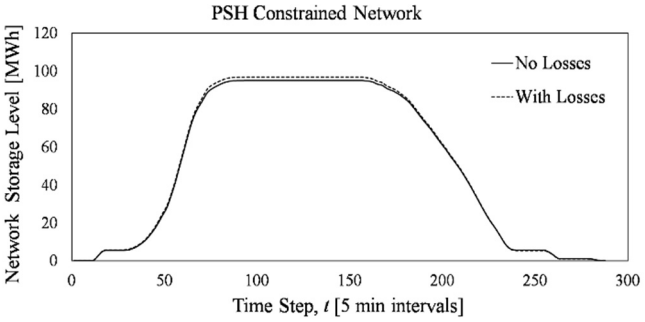


Figure 5-3. PSH Storage in a congested network

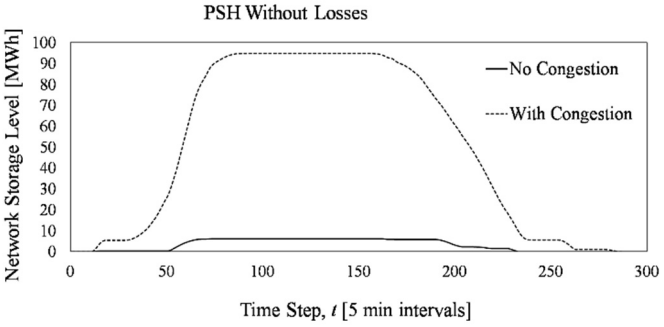


Figure 5-4. PSH storage with and without congestion

In Figure 5-3 and Figure 5-4, we see that losses take a backseat as the driver of capacity allocation in congested networks, creating no significant shifts in storage operations. This is best illustrated in Figure 5-3 by looking at how the storage level curve with losses has negligible differences to the based case curve. introduced. However, when congestion was introduced on its own, we observe a drastic impact on PSH storage as shown by the dramatic increase in system wide storage level in Figure 5-4. Thus, while it may be important to consider transmission losses in a free-flowing network, in a congested they will not significantly impact how a storage system should be integrated.

5.1.7 Results for ESS Investment in an Unconstrained Network

While the previous cases worked on the premise that all the available storage capacity is allocated and free of charge, the next two cases will include the cost of installation as well. In every one of the four scenarios investigated for allocation, FES was the only storage technology worth investing in. This is because of its flexibility in deployment, and ability to stabilize load. However, no significant changes in capacity investment were observed when losses, congestion, or both forms of blockage were introduced. We can conclude that network inhibitors did not increase the marginal value of storage capacity, making the investment costs too high to bare for large-scale expansion at the base level of renewable penetration.

To investigate the value-add that storage can provide as a load stabilizing technology, the net demand profile was shifted by introducing different degrees of wind generation. Since the erratic generation pattern of wind was likely pushing all investment into FES, we wanted to understand what would happen with different magnitudes of wind generation. The NREL Western Wind Sources Data Set [136] was applied to the model at 25%, 50%, 100%, 200%, 250% and 300% of the base values used in Sections 5.1.5 and 5.1.6. The Standard Value of wind production will be used to refer to the case of 100% of production provided in the NREL Western Wind Data Set.

For the case of an unconstrainted network at or below standard wind production, FES was the only storage technology invested in. In fact, when wind generation was reduced to the lower boundary of 25% of standard production, the total invested capacity of FES only dropped by 5% as indicated by the capacity investment values in Table 5-3.

This implies that the need for FES capacity may primarily be to provide general load balance as opposed to just dealing with fluctuations from volume of

wind generation. Despite this, when wind production is increased in the network, a significant change in investment strategy is observed.

At a 2.5-fold increase, investment in FES capacity increases by 41% and with an additional 103 MWh of CAES being invested as shown in the second to last column of Table 5-3. This is the first instance where the model chose to invest in CAES across all scenarios. Recall that in the allocation cases, the low efficiency of this technology seemed to make it an unfavorable choice. Similarly, at low levels of wind generation, CAES was not advantageous to install.

However, when wind production rose (and thus overall energy produced in the network increased) there grew a need for large-scale energy reservoirs. Table 5-3 indicates this pattern in the first case, right-most column, by the high volume of capacity investment (i.e., darkest green cells). The cost of installing new technologies is a function of the discharge duration as well as a capacity cost. Technologies with longer discharge durations have a lower cost per MWh. Looking back to Table 5-1, we see that likely this caused the model to favor CAES despite its inefficiency. In the Standard Wind case, a significantly more expensive technology like FES is favorable because of its ability to smooth out the energy balance. In the high production scenarios, the sheer increase in volume of energy seems to be the most significant driver of investment decisions. The model chooses to find a way to use all the power in the network anyway it can as a priority. It should be noted that in the 300% case, no energy was produced from thermal generation.

Table 5-3. Total invested capacity for each technology in kWh

Case Study	Tech	Wind production					
		x 0.25	x 0.50	x 1.0	x 2.0	x 2.5	x 3.0
Congestion = NO Losses = NO	PSH	0	0	0	0	9	206
	CAES	0	0	0	0	103	535
	LION	0	0	0	0	0	55
	FES	200	200	211	237	298	0
Congestion = NO Losses = YES	PSH	0	0	0	0	19	211
	CAES	0	1	0	0	123	538
	LION	0	0	0	0	0	57
	FES	204	204	217	257	312	0
Congestion = YES Losses = NO	PSH	0	0	0	0	49	398
	CAES	1	1	0	0	209	965
	LION	0	0	0	0	14	116
	FES	188	266	212	363	309	16
Congestion = YES Losses = YES	PSH	0	0	0	0	53	373
	CAES	0	1	0	0	204	818
	LION	0	0	0	0	13	117
	FES	184	262	221	361	326	17

Figure 5-5 can be used to get an idea of the distribution patterns for the investment models using the FES investment results in the base case wind production scenario. First, transmission congestion has more impact on the distribution of investment than the transmission losses. However, if transmission losses are not considered then the FES capacity is invested in evenly throughout the network, making it impossible to differentiate among the

nodes for investment. This situation changes when losses are introduced, for instance, there are significant increases at nodes 3, 4 and 9, which comprise the three largest demand nodes. There is also a significant increase in storage capacity at nodes 7 and 8. These two nodes house major wind generation and losses along them must be incurred to reach the other parts of the network. By storing power at the generation source, a system operator gains some flexibility in where he can distribute stored power when it is needed. Generally, in the unconstrained case, losses are a major driving force in ESS investment throughout the network. This causes the system operator to change his investment strategy to best serve locations where demand is highest and thus power is needed most. Keeping storage capacity nearby allows the system operator to prevent additional losses that would occur if power had to be first transmitted from generator nodes to storage nodes and then back to the demand centers.

Congestion = NO & Losses = NO

	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	n11	n12	n13	n14
x 3.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x 2.50	21	21	21	21	21	21	21	21	21	21	21	21	21	21
x 2.00	17	17	17	17	17	17	17	17	17	17	17	17	17	17
x 1.00	15	15	15	15	15	15	15	15	15	15	15	15	15	15
x 0.50	14	14	14	14	14	14	14	14	14	14	14	14	14	14
x 0.25	14	14	14	14	14	14	14	14	14	14	14	14	14	14

Congestion = NO & Losses = YES

	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	n11	n12	n13	n14
x 3.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x 2.50	30	18	49	30	13	42	19	19	19	10	6	15	26	16
x 2.00	13	20	39	26	9	37	16	16	16	8	5	14	23	16
x 1.00	10	12	26	25	3	29	14	14	14	9	5	12	23	21
x 0.50	9	11	19	26	2	20	16	16	13	10	4	11	23	24
x 0.25	9	11	18	23	4	15	17	17	14	11	4	12	23	26

Congestion = YES & Losses = NO

	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	n11	n12	n13	n14
x 3.00	0	0	13	0	0	0	0	0	0	0	0	0	0	0
x 2.50	2	2	189	17	5	7	14	14	12	11	9	8	8	10
x 2.00	4	25	281	10	1	2	7	7	6	5	4	2	3	4
x 1.00	0	4	137	34	0	0	14	14	6	2	0	0	0	1
x 0.50	0	1	66	116	0	0	34	34	12	3	0	0	0	0
x 0.25	0	1	50	81	0	0	24	24	6	1	0	0	0	0

Congestion = YES & Losses = YES

	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	n11	n12	n13	n14
x 3.00	0	0	13	1	0	0	0	0	0	0	0	0	0	0
x 2.50	5	5	198	15	6	30	10	10	10	5	3	8	12	9
x 2.00	1	24	275	6	1	15	6	6	6	3	1	4	6	6
x 1.00	0	3	135	36	0	3	6	6	1	2	0	6	12	10
x 0.50	0	1	66	117	0	5	20	20	0	2	1	5	13	13
x 0.25	0	1	50	81	0	1	10	10	0	1	0	5	12	13

Figure 5-5. FES invested capacity per node

5.1.8 Results for ESS Investment in a Constrained Network

Congestion was introduced identically to the allocation case, as follows: $TC_{12}^{MAX} = 80MW, TC_{15}^{MAX} = 40MW, TC_{23}^{MAX} = TC_{42}^{MAX} = 30MW$. As was the case in the allocation model, congestion had a more significant impact on the investment strategy than did losses. In the free-flowing case, most of the system wide storage capacity was sited at the bus with the greatest load—bus 3. When wind generation increases over twice the standard rate, FES is replaced with CAES. The introduction of losses alone only led to a slight increase in system storage capacity while maintaining the same technology selection shown in Figure 5-7. This held true even for the highest wind generation scenarios. When congestion was introduced into the network, system-wide storage capacity nearly doubled. Most of this change can be attributed to an 80% increase in CAES capacity investment.

Much like the free-flowing case, the largest share of storage capacity, specifically CAES, was located at bus 3. Looking at Table 5-3, we observe that as wind generation increased from 250% to 300%, this pattern was exemplified. It manifested itself in a nearly 5-fold increase in CAES, 5.5 times increase in PSH and a doubling of LION at bus 3. These results are reflected in the two right-most columns of Table 5-3. This was significantly higher than the increases observed when introducing losses alone. The model seemed to generally follow this pattern as wind generation in the system increased.

Adding both transmission constraints and losses into the model dampened these effects slightly. While there were significant changes in storage capacity investment from observing losses alone in the model, the network appears to be too constrained to invest in quite as much storage. Inhibiting power flow makes it less economic to have storage located only at demand and generation hubs. Since losses must be incurred when moving throughout the network, it makes less sense both logistically and economically to rely on stored power. For this reason, a decrease in overall storage capacity of 10% is observed.

5.1.9 Analysis and Discussion

To better understand the results of the case studies, we will explore the implications of the observations made earlier. For the various allocation scenarios, we find that round-trip efficiency governs the way the storage capacity is sited, and therefore the technology with the lowest efficiency, CAES was never chosen. When investigating the effects of losses and congestion, we find that congestion is a more forceful driver in capacity allocation than are losses. Both network inhibitors push capacity siting towards nodes with high demand and high wind generation, with congestion doing so to a much greater degree. Introducing losses to a congested network results in no change in capacity siting or in technology deployment. It is therefore imperative for a system operator to consider network congestion above losses as an unforeseen cost when planning grid-scale storage integration. As for transmission losses alone, ignoring them may lead to some initial saving but long term will require network upgrades.

The second set of cases focused on storage capacity investment. When running the same scenarios as the allocation model, we find that the model output did not vary significantly with the introduction of congestion or losses. The model always favors investment in FES storage since the marginal value of storage capacity is tied to its load stabilizing capabilities. To study different degrees of load variability, the cases were rerun with varying levels of renewable penetration.

When wind production is scaled down no substantial changes to the FES investment strategy are observed, affirming that this technology's primary use is for system load-balance needed irrespective of non-dispatchable generation. However, when production is increased, the investment strategy changes greatly, buying mostly into CAES capacity. Since the model does not allow for any spillage, all wind power generated must be consumed to meet demand, lost via transmission or stored. We find that losses do not significantly change the ESS investment mix at any level of renewable penetration, however, they can help to differentiate among the nodes for provide a spatial distribution for ESS investments. On the other hand, congestion leads to nearly double system-wide storage capacity for the highest wind production case. This is in line with the observations made when looking at the allocation model—congestion plays a more important factor in planning large-scale energy storage integration than losses do.

Looking at the objective function output for these cases we can reaffirm this observation. Note that the demand in every case was held constant. The base case objective function value (equivalently system operating cost) was

$841,891. For the allocation model, introducing losses, only causes a 2.4% increase in operating cost, while congestion results in a 13.5% increase. When adding losses to the constrained network, we only see an additional increase of 2% to the objective function value.

The objective function in the base case for the investment model was $836,343 (0.7%). A summary of the per-unit change in system operating costs for the set of investment studies can be seen in Table 5-4. It can be seen that with standard wind generation, losses again only add 2.4% to the objective function value while congestion increases costs by a factor of 12.9%. As was shown in the allocation model, the addition of losses to a congested system only increases the overall operating cost by 2%. Moving horizontally across the table we observe that operational cost has an inverse relationship to the amount of wind generation available. Since demand is held constant in all cases, as free wind generation decreases, thermal generation must be deployed to compensate. This cost increase becomes larger in a constrained network.

Table 5-4. Change in objective value [p.u.] for investment case studies

Case Study	Wind Production					
	x 0.25	**x 0.50**	**x 1.0**	**x 2.0**	**x 2.5**	**x 3.0**
Congestion = NO Losses = NO	1.645	1.419	1.000	0.290	0.031	0.000
Congestion = NO Losses = YES	1.681	1.451	1.024	0.301	0.037	0.000
Congestion = YES Losses = NO	1.854	1.598	1.129	0.307	0.035	0.000
Congestion = YES Losses = YES	1.880	1.622	1.149	0.320	0.041	0.000

5.2 System States and Representative Periods Comparison

The previous Section has shown the impact of transmission congestion and losses in the ESS investment decisions. The results were focused on STESS investment, while the LTESS was not considered in the case studies. In this section, we address the ESS investment considering at the same time STESS and LTESS. This is especially important in power systems with seasonal hydro reservoirs. In Section 3.2, the models based on the SS and RP approaches were compared from the operational point of view. Here the classic SS and RP approaches and their enhanced versions (SS-RFM and RP-TM&CI) are tested considering ESS investment decisions. The procedure to compare the results is the same as the one used to compare the operating results, see Section 3.2.1. Therefore, the HM model is used as a benchmark in the comparison. All formulations are detailed in Appendix A.

5.2.1 Case Study and Results

For the investment analysis, we use the same case as in Section 3.2.5. However, only the investment results are shown because the trend is similar to that of the operational results (e.g. production, number of start-ups, prices), see Section 3.2.6.

We consider the possibility of investment in BESS technology. Unlike the operation case study in Section 3.2.5, BESS initial capacity is not predefined. We consider an investment cost of 20 [€/kW] for BESS according to the report "*Technology development roadmap towards 2030*" [137] and a maximum energy to power ratio of 4 hours. Table 5-5 shows objective function error and investment error for each scenario using the HM model results as a reference. All four models underestimate the objective function, especially when there is a high share of VRES (scenario 4). However, the range of the error values remains similar to those shown in Figure 3-11 in the operation analysis in Section 3.2.6. As for the investment error, the RP-TM&CI model offers the best approximation. This is because it is the model that most accurately estimates energy prices and energy production of each technology, as it was shown in Section 3.2.6. Both the SS-RFM and RP-TM&CI models, the original contributions of this thesis, represent significant improvements on their former versions of SS and RP.

Table 5-5. Investment result error [%] per scenario

Result	Scenario	SS	SS-RFM	RP	RP-TM&CI
Objective Function Error	V1	0.5	0.1	0.8	0.1
	V2	1.2	1.0	4.4	0.7
	V3	0.5	0.4	4.8	5.4
	V4	6.4	6.5	1.8	5.6
BESS Investment Error	V1	72.4	52.0	38.3	-10.3
	V2	35.4	32.9	22.2	-8.3
	V3	57.1	49.8	32.3	-2.5
	V4	34.4	28.8	31.7	3.9

Figure 5-6 shows BESS investment obtained with all the models for each scenario, and the share of VRES (i.e., wind and solar productions). As expected, BESS investment increases when the VRES share increases in the power system. The SS model and the SS-RFM model underestimate the investment by the greatest amount due to their main drawback, which is that they do not fully guarantee that the energy stored in the batteries is lower than the capacity of the batteries. This means that they permit energy to be stored beyond what investment has paid for, and therefore require less investment to achieve the same results as the RP model and the RP-TM&CI model.

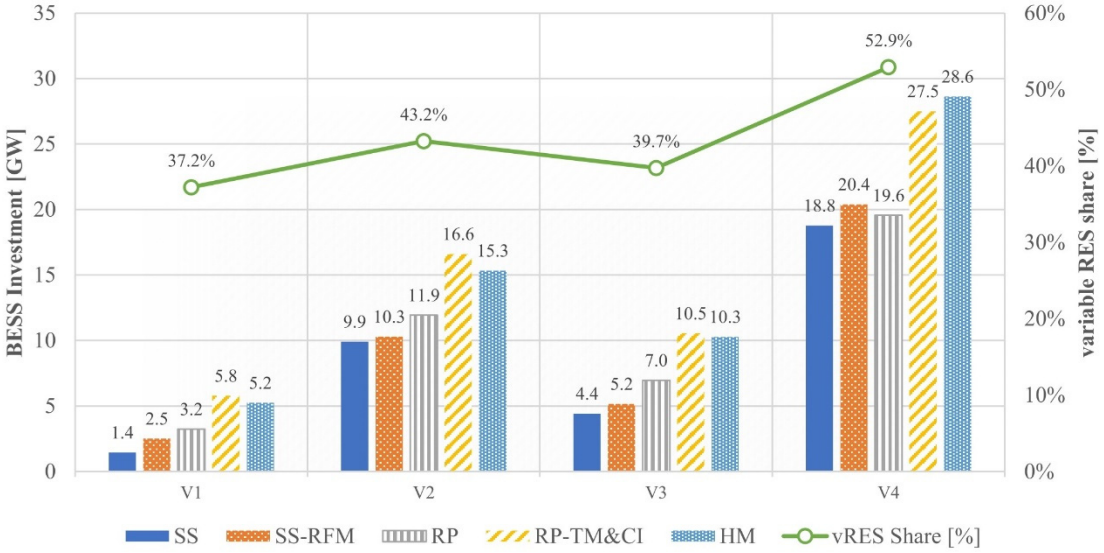


Figure 5-6. BESS investment and VRES share for each scenario

5.2.2 ESS Investment and VRES Curtailment

There is an intrinsic relationship between ESS investment and VRES curtailment. For instance, Figure 5-7 shows the VRES curtailment as a percentage of the total available VRES for each scenario.

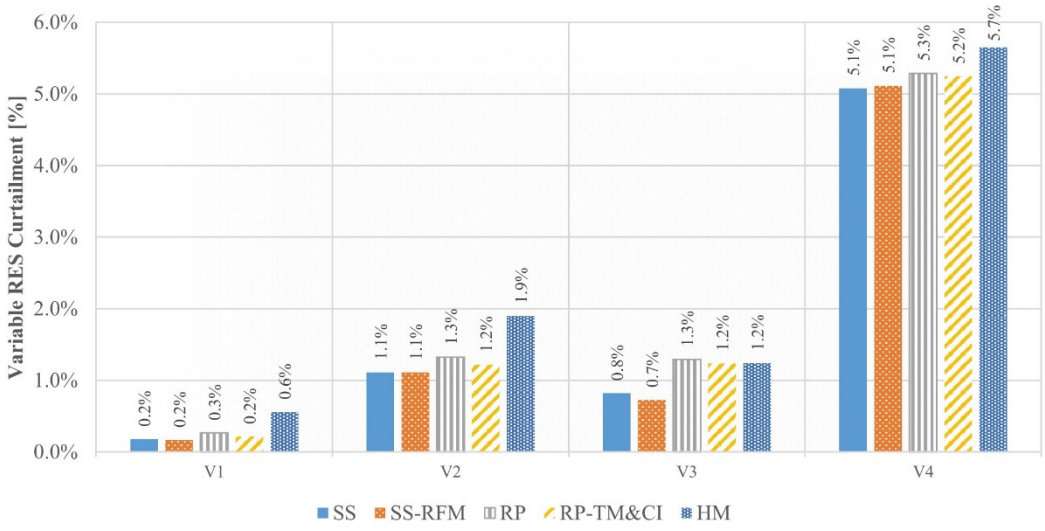


Figure 5-7. Variable VRES curtailment for each tested model

The amount of curtailment determined by all models underestimates the reference values from the hourly model. While a portion of the under-investment in storage shown in Figure 5-6 is due to the inaccuracies in the way storage is represented in each model, some of the underinvestment may also come from the models' underestimation of VRES curtailments. This is based on the tight connection between VRES curtailment and storage needs, as shown in [138]. Models such as ReEDS and RPM use exogenous estimations to relate these two aspects in systems with high share of VRES. However, the models proposed in this thesis determine this relationship endogenously, because the

renewable production is a variable in the optimization problem. Improvements in the clustering process could be performed to improve this relationship; however, further research is needed to verify this hypothesis.

5.3 Main Takeaways in this Chapter

Through this chapter we have studied optimization models for ESS investment in medium- and long-term planning. The analyses are divided in two main sections. First, we have investigated the effects of transmission losses, constraints and increased renewable energy penetration on planning ESS allocation and investment, using a DC Optimal Power Flow model with a linearized approximation for ohmic losses. Second, we have compared the system-states type and representative-periods type methods when short-term energy storage and seasonal storage are included in the optimization models for ESS investment decisions. The analysis and results can be summarized in the following conclusions:

- *Transmission losses in the case study had a relatively small effect on storage siting and investment decisions, while congestion significantly impacted the model output. Congestion drove higher storage capacity siting with more of it located close to demand centers. When both forms of network constraints were introduced simultaneously, losses did not drastically change the results.*
- *In addition, the case studies showed that faster ramping technologies such as flywheels were favored for increased renewable penetration for their load stabilizing capabilities. However, after a certain threshold of renewable capacity was reached the investment model favored larger-scale ESS to better manage excess energy supply and avoid spillage.*
- *Nevertheless, if there is no network congestion then the transmission losses play an important role to differentiate among nodes to distribute the ESS investment.*
- *When both STESS and LTESS are considered for investment decisions, then SS and RP approaches are also useful to determine the optimal investment decisions. As in the operational results in Section 3.2.6, the enhanced version of RP (i.e., RP-TM&CI model) is the most accurate of the four approximate models. Therefore, including chronological information among representative periods may be an efficient way to include short-term variations in longer-term planning models that involve energy storage investment.*
- *The system states and representative periods models, as well as their enhanced versions, underestimate the ESS investment. Based on the tight connection between VRES curtailment and storage needs, the results show that the representative periods models (RP and RP-TM&CI) have less error (both ESS investment and renewable curtailment) than the system states models (SS and SS-RFM).*

6 POWER-BASED MODEL FOR FLEXIBILITY REQUIREMENTS

Contents

Previous chapters have focused on both STESS and LTESS operational and investment decisions. However, flexibility requirements due to the integration of large amounts of renewable energy sources have been disregarded. For instance, previously proposed optimization models do not fully consider detailed unit commitment constraints (e.g., minimum up/down time, ramping constraints, startup/shutdown trajectories). In addition, these models use an energy-based unit commitment formulation, which overestimate the flexibility of the system. Power-based UC models overcome these problems by correctly modeling ramping constraints and operating reserves. This chapter proposes a power-based GEP-UC model that improves the classic energy-based models by representing more accurately the flexibility requirements of power systems (i.e., reserve decisions and ramping constraints), as well as a real-time validation stage (e.g., 5-min simulation) in order to evaluate the quality of investment and operational decisions obtained with the model. Moreover, we also propose

a semi-relaxed version of the power-based GEP-UC model, which aims at reducing the computational burden without losing accuracy in the results.

The results show that the power-based model uses the installed investments more effectively than the energy-based model because it represents flexibility capabilities and system requirements more accurately. For instance, for the analyzed case study, the power-based model obtains less investment (6-12%) and yet it uses this investment more efficiently because operating cost is also lower (2-8%) in a real-time validation. The semi-relaxed power-based model is at least 10 times faster than its full-integer version and without significantly losing accuracy in the results (less than 0.2% error).

The analysis and results in this section were published in Article VII [139] *as working paper, which is under review at the time of the publication of this thesis.*

6.1 Flexibility Requirements and Modeling Options

In the last section of Chapter 2, we have introduced the basic concepts of the Unit Commitment (UC) problem and its important role in long-term planning problems, e.g., Generation Expansion Planning (GEP), in order to consider the increased need of operational flexibility due to VRES integration. Startup and shutdown processes, ramping constraints, minimum up/down time, operating reserve are among the main constraints to consider flexibility requirements. We have also mentioned that recent studies [31], [32], [70] have shown that energy-based UC models cannot capture variability on demand and VRES, and even assuming that they could capture it, they cannot deliver the flexibility that they promise, that is, they overestimate the flexibility of the system. In addition, we have discussed that power-based models [70], [71] overcome these problems by better exploiting the system flexibility [31], by allowing the correct modeling of ramping constraints and operating reserves [32], [70] in order to deliver the expected and actual flexibility from the generation resources. This is possible because a power-based model has a clear distinction between power and energy in its core formulation. Demand and generation are modeled as hourly piecewise-linear functions representing their instantaneous power trajectories. The schedule of a generating unit output is no longer an energy stepwise function, but a smoother piece-wise power function. The following section shows this distinction through a detailed comparison among the different constraints in both formulations.

6.2 Energy-based and Power-based Models

This section presents the objective function and set of constraints for the energy- and the power-based GEP-UC models. These constraints include investment decisions for different generation technologies: thermal generation, ESS, and VRES. In addition, operational decisions are considered using a clustered UC formulation (i.e., aggregating similar generating units into one group or cluster). This type of aggregation is commonly applied in long-term planning models [60], [140]. As part of this thesis, Article V [141] discusses the implications of the Clustered Unit Commitment (CUC) formulations to

represent flexibility requirements, and it proposes a set of constraints to improve the CUC constraints such as ramping, reserve, and startup/shutdown. For the sake of simplicity, this set of constraints is not presented in detail in this section. However, the reader is referred to Article V [141] at the end of this thesis for more information about the proposed CUC constraints.

For the sake of briefness, this section only highlights the main differences between the energy- and power-based models. Nevertheless, the complete list of sets, variables, parameters, and equations for both models is detailed in Appendix D.

6.2.1 Objective Function

The GEP seeks to minimize the investment costs plus the expected value of operating costs: production cost, up/down reserve cost, CO2 emission cost, no-load cost, shutdown cost, and startup cost.

$$\min_{\Psi} \underbrace{\sum_{j\in\mathcal{J}} C_j^I x_j}_{\substack{investment\\cost}} + \sum_{\omega\in\Omega}\pi_\omega \sum_{t\in\mathcal{T}} \left\{ \sum_{j\in\mathcal{J}} \left[\underbrace{C_j^{LV}\hat{e}_{\omega jt}}_{\substack{energy\\production\ cost}} + \underbrace{C_j^{R+}r_{\omega jt}^{+} + C_j^{R-}r_{\omega jt}^{-}}_{reserve\ cost} \right] \right.$$
$$+ \sum_{g\in\mathcal{G}} \left[\underbrace{C_g^{EM}\hat{e}_{\omega gt}}_{\substack{CO2\ emmision\\cost}} + \underbrace{C_g^{NL}u_{\omega gt}}_{\substack{no-load\\cost}} + \underbrace{C_g^{SD}z_{\omega gt}}_{\substack{shutdown\\cost}} \right. \tag{6-1}$$
$$\left. \left. + \underbrace{\sum_{k\in\mathcal{K}_g} C_{gk}^{SU}\delta_{\omega gkt}}_{startup\ cost} \right] \right\}$$

Equation (6-1) shows the objective function, where parameters: π_ω is probability of scenario ω, C_j^I is the investment cost per installed MW of technology j, C_j^{LV} is the linear variable energy production cost \$/MWh, C_j^{R+} and C_j^{R-} are up/down reserve cost \$/MW, C_g^{EM} is the CO2 emission cost \$/MWh, C_g^{NL} is the no-load cost per hour, C_{gk}^{SU} is the startup cost, and C_g^{SD} is the shutdown cost. In addition, the variables: x_j is the investment decisions per technology, $\hat{e}_{\omega jt}$ is the total energy output at time t for technology j in scenario ω, $r_{\omega jt}^{+}$ and $r_{\omega jt}^{-}$ are up/down reserves, $u_{\omega gt}$ is the unit commitment at time t for thermal technology g in scenario ω, $z_{\omega gt}$ is the shutdown decision variable, and $\delta_{\omega gkt}$ is the startup decision variable depending on the startup segment k.

In the energy-based model, the set of decision variables Ψ in (6-1) includes previous variables plus: energy output above minimum output (e), charged energy in a storage technology ($\hat{c}$), energy storage level (ϕ), and binary decision for charging/discharging logic (γ). In the power-based model, the set of decision variables Ψ additionally includes power related variables, such as total power output ($\hat{p}$), power output above minimum output (p), and charged power for a storage technology (c). Table 6-1 summarizes the decision variables in set Ψ depending on the model.

Table 6-1. Set of decision variables Ψ for each model

Model	Decision variables
Energy-based	$\Psi = \{x, e, \hat{e}, \hat{c}, r^+, r^-, u, y, z, \delta, \gamma, \phi\}$
Power-based	$\Psi = \{x, p, \hat{p}, \hat{e}, c, \hat{c}, r^+, r^-, u, y, z, \delta, \gamma, \phi\}$

In order to obtain the total energy output in the power-based model, we use equation (6-2) where energy is represented as a function of total power output at the end of time step t and $t-1$. The same concept is applied for the total charged energy in energy storage technology s.

$$\hat{e}_{\omega jt} = \frac{\hat{p}_{\omega jt} + \hat{p}_{\omega j,t-1}}{2} \tag{6-2}$$

$$\hat{c}_{\omega st} = \frac{c_{\omega st} + c_{\omega s,t-1}}{2} \tag{6-3}$$

6.2.2 System-wide Constraints

The system-wide constraints are represented by demand balance, transmission line limits, and up/down reserve requirements. Table 6-2 shows these constraints depending on which model is used. Here, $D^E_{\omega bt}$ and $D^P_{\omega bt}$ are correspondently the energy and power values of demand, $\overline{F}$ is the maximum power flow on a transmission line, Γ^J_{lj} and Γ_{lb} are the shift factors, also $R^+_{\omega t}$ and $R^-_{\omega t}$ are the up/down reserve requirements of the system. Notice that up/down reserve requirement constraints are the same for both models because reserve requirements are expressed in terms of power, as they represent an instantaneous need in the power system. Nevertheless, transmission line limits and hourly demand balance constraints are different depending on the type of model. In the power-based model, demand ($D^P_{\omega bt}$) and technology production/charging ($\hat{p}_{\omega jt}$, $c_{\omega st}$) are modeled as hourly piecewise-linear functions representing their instantaneous power values at the end of each hour. In contrast, demand ($D^E_{\omega bt}$) and technology production/charging ($\hat{e}_{\omega jt}$, $\hat{c}_{\omega jt}$) are modeled as an energy stepwise function in the energy-based model. This situation was illustrated in the background section (Figure 2-10) using $D^E_{\omega bt}$ and $D^P_{\omega bt}$.

Table 6-2. System-wide constraints

Constraint	Energy-based	Power-based
Demand balance	$\sum_{j\in\mathcal{J}} \hat{e}_{\omega jt} - \sum_{s\in\mathcal{S}} \hat{c}_{\omega st} = \sum_{b\in\mathcal{B}^D} D^E_{\omega bt}$ (6-4)	$\sum_{j\in\mathcal{J}} \hat{p}_{\omega jt} - \sum_{s\in\mathcal{S}} c_{\omega st} = \sum_{b\in\mathcal{B}^D} D^P_{\omega bt}$ (6-5)
Transmission line limits	$-\overline{F}_l \le \sum_{j\in\mathcal{J}} \Gamma^J_{lj}\hat{e}_{\omega jt} - \sum_{s\in\mathcal{S}} \Gamma^S_{ls} c_{\omega st} - \sum_{b\in\mathcal{B}^D} \Gamma_{lb} D^E_{\omega bt} \le \overline{F}_l$ (6-6)	$-\overline{F}_l \le \sum_{j\in\mathcal{J}} \Gamma^J_{lj}\hat{p}_{\omega jt} - \sum_{s\in\mathcal{S}} \Gamma^S_{ls} c_{\omega st} - \sum_{b\in\mathcal{B}^D} \Gamma_{lb} D^P_{\omega bt} \le \overline{F}_l$ (6-7)
Up Reserve requirement	$\sum_{j\in\mathcal{J}} r^+_{\omega jt} \ge R^+_{\omega t}$ (6-8)	
Down Reserve requirement	$\sum_{j\in\mathcal{J}} r^-_{\omega jt} \ge R^-_{\omega t}$ (6-9)	

6.2.3 Investment Constraints

Investment decision variables are commonly expressed as lumped values in GEP models. This follows the fact that maximum capacities values frequently depend on efficient size/cost ratios and standardized sizes of their components. Therefore, the investment variable x_j is an integer decision variable in both models.

The relationship between operational and investment decisions for each technology type is guaranteed using the constraints in Table 6-3. The thermal investment constraint is the same for both models since it depends on integer variables associated to the number of units connected within the thermal technology cluster and the investment decisions (none of these variables are related to energy or power concept). However, ESS and VRES investment constraints change depending on the model since their operational decision variables change depending on whether the model is using power-related variables or energy-related variables.

X_j^0 is a parameter to define the initial capacity for technology j; [# units] for g, and [MW] for s and v.

Table 6-3. Investment constraints

Constraint	Energy-based	Power-based
Thermal technologies	$u_{\omega gt} \leq X_g^0 + x_g$ (6-10)	
ESS technologies	$\hat{e}_{\omega st} - \hat{c}_{\omega st} + r_{\omega gt}^{+} \leq X_s^0 + x_s$ (6-11)	$\hat{p}_{\omega st} - c_{\omega st} + r_{\omega gt}^{+} \leq X_s^0 + x_s$ (6-12)
	$\hat{e}_{\omega st} - \hat{c}_{\omega st} - r_{\omega gt}^{-} \geq -(X_s^0 + x_s)$ (6-13)	$\hat{p}_{\omega st} - c_{\omega st} - r_{\omega gt}^{-} \geq -(X_s^0 + x_s)$ (6-14)
VRES technologies	$\hat{e}_{\omega vt} \leq V_{\omega vt}^{E}(X_v^0 + x_v)$ (6-15)	$\hat{p}_{\omega vt} \leq V_{\omega vt}^{P}(X_v^0 + x_v)$ (6-16)

As they are written here, constraints for ESS technologies might lead to a solution in which the ESS charge and discharge at the same time period t. This is physically impossible in an ESS; however, it could be the least cost solution considering that we have reserve decision variables. For instance, let us consider a feasible solution where the ESS is charging and discharging using full capacity, these set of constraints also allow the ESS to provide upward and downward reserve up to its maximum capacity, which is not physically possible. Table 6-4 shows a set of constraints in order to avoid this situation in both models, which is based on [142]. $\overline{X}_s$ is a parameter that represents the maximum investment limit for technology j, whereas $\gamma_{\omega st}$ is a binary variable that ensures either charging or discharging mode in the ESS. Therefore, reserve decisions for ESS now are different depending on whether the ESS is charging or discharging. Figure 6-1 shows an ESS providing reserves from different operation points within its limits thanks to the set of constraints in Table 6-3 combined with constraints in Table 6-4.

Table 6-4. ESS charge/discharge logic constraints

Constraint	Energy-based	Power-based
ESS logic 1	$\hat{c}_{\omega st} \le (1-\gamma_{\omega st}) \cdot (X_s^0 + \overline{X}_s)$ (6-17)	$c_{\omega st} \le (1-\gamma_{\omega st}) \cdot (X_s^0 + \overline{X}_s)$ (6-18)
ESS logic 2	$\hat{e}_{\omega st} \le \gamma_{\omega st} \cdot (X_s^0 + \overline{X}_s)$ (6-19)	$\hat{p}_{\omega st} \le \gamma_{\omega st} \cdot (X_s^0 + \overline{X}_s)$ (6-20)

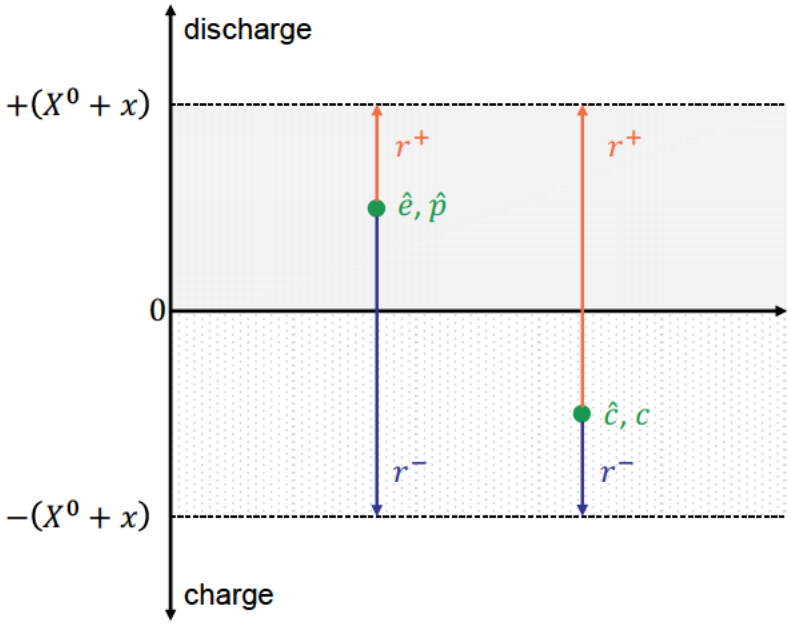


Figure 6-1. ESS providing reserves from different operation points

6.2.4 UC Constraints

The UC constraints relate the operational integer decision variables among them. Therefore, these constraints include commitment/ startup/ shutdown logic (6-21), minimum up/down times (6-22)-(6-23), startup type selection (6-24)-(6-25) (e.g., hot, warm, and cold startup). Where, TU_g and TD_g are the minimum up/down time correspondently, T_{gk}^{SU} is the time interval limit of startup segment k.

Since these constraints are independent of energy or power concepts, they are the same for both models.

$$u_{\omega gt} - u_{\omega g,t-1} = y_{\omega gt} - z_{\omega gt} \tag{6-21}$$

$$\sum_{i=t-TU_g+1}^{t} y_{\omega gi} \le u_{\omega gt} \tag{6-22}$$

$$\sum_{i=t-TD_g+1}^{t} z_{\omega gi} \le \left(X_g^0 + x_g\right) - u_{\omega gt} \tag{6-23}$$

$$\delta_{\omega gkt} \le \sum_{i=T_{gk}^{SU}}^{T_{g,k+1}^{SU}-1} z_{\omega g,t-i} \tag{6-24}$$

$$\sum_{k \in \mathcal{K}_g} \delta_{\omega g k t} = y_{\omega g t} \tag{6-25}$$

6.2.5 Thermal Generation Technology Constraints

These constraints correspond to the maximum and minimum output limit of thermal generation technologies.

Table 6-5. Thermal generation technology constraints

Constraint	Energy-based	Power-based
Maximum output limit	$e_{\omega g t} + r^{+}_{\omega g t} \le (\overline{P}_g - \underline{P}_g) u_{\omega g t} - (\overline{P}_g - SD_g) z_{\omega g, t+1} - \max(SD_g - SU_g, 0) y_{\omega g t}$ $\forall g \in \mathcal{G}^1$ (6-26)	$p_{\omega g t} + r^{+}_{\omega g t} \le (\overline{P}_g - \underline{P}_g) u_{\omega g t} - (\overline{P}_g - SD_g) z_{\omega g, t+1} + (SU_g - \underline{P}_g) y_{\omega g, t+1}$ (6-27)
	$e_{\omega g t} + r^{+}_{\omega g t} \le (\overline{P}_g - \underline{P}_g) u_{\omega g t} - (\overline{P}_g - SU_g) y_{\omega g, t} - \max(SU_g - SD_g, 0) z_{\omega g, t+1}$ $\forall g \in \mathcal{G}^1$ (6-28)	
	$e_{\omega g t} + r^{+}_{\omega g t} \le (\overline{P}_g - \underline{P}_g) u_{\omega g t} - (\overline{P}_g - SU_g) y_{\omega g, t} - (\overline{P}_g - SD_g) z_{\omega g, t+1}$ $\forall \omega, g \notin \mathcal{G}^1$ (6-29)	
Minimum output limit	$e_{\omega g t} - r^{-}_{\omega g t} \ge 0$ (6-30)	$p_{\omega g t} - r^{-}_{\omega g t} \ge 0$ (6-31)

Constraints (6-26),(6-28)–(6-30) are associated to the energy-based model, while (6-27),(6-31) are associated to the power-based model. Where $\mathcal{G}^1$ is defined as the thermal technologies in $\mathcal{G}$ with $TU_g = 1$, $\overline{P}_g$ and $\underline{P}_g$ are the maximum/minimum power output, and SU_g and SD_g are the startup/shutdown capability.

Figure 6-2 and Figure 6-3 show the operation of thermal generation technologies for each model. On the one hand, Figure 6-2 shows that the energy-based model considers energy as the direct output of thermal generation technologies, as well as zero production below $\underline{P}_g$. On the other hand, notice that Figure 6-3 shows the power output of generating units, while the energy is then obtained from the power profile using (6-2). Moreover, the resulting energy using (6-27) takes values lower than $\underline{P}_g$ during the startup/shutdown, different from the traditional energy-based UC in Figure 6-2. Section 6.2.6 shows in more detail these differences during the start/shutdown process in both models.

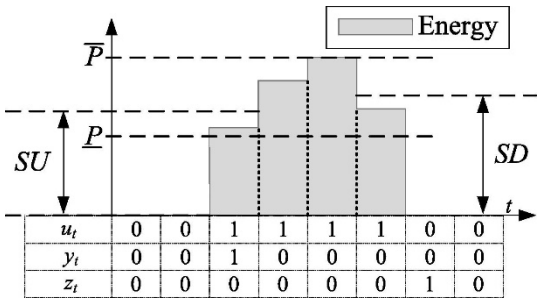


Figure 6-2. Energy-based representation of operation [31]

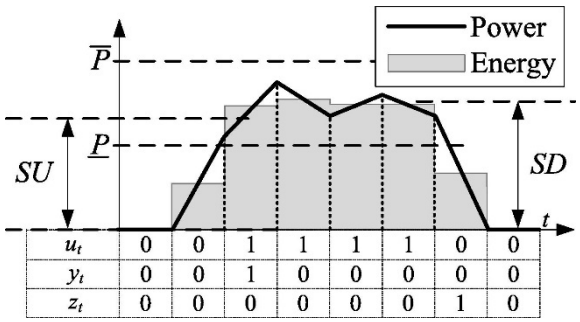


Figure 6-3. Power-based representation of operation [31]

6.2.6 Total Thermal Generation Technology Output

The previous section has explained how the energy-based model considers energy as the direct output of generation units. This consideration leads to generation schedules in which the units produce either above $\underline{P}_g$, if the unit is on-line, or zero, if the unit is off-line. Startup (SU) and shutdown (SD) capabilities limit the maximum output during the startup/shutdown process. This consideration may be valid for quick-start units (i.e., units that can startup in less than one hour). However, it is not valid for slow-start units that commonly have startup durations greater than one hour, see Figure 6-4. Therefore, traditional energy-based formulations ignore the inherent startup and shutdown trajectories of thermal generation. Figure 6-4 shows the energy-based formulation including the SU/SD trajectories proposed in [143] in order to overcome this drawback. Authors in [31], [70] have shown the relevance of the SU and SD processes when they are included in the scheduling optimization. Thus, if the traditional energy-based formulation is used, constraint (6-32) defines the total production output of thermal generation technologies, while if SU/SD trajectories are considered in the energy-based formulation then (6-32) is replaced by (6-33).

In the power-based model, the total power output constraint is different depending on whether it is a quick- or slow-start unit. As mentioned before, quick-start technologies $\mathcal{G}^F$ are thermal generators that can startup/shutdown within one hour (i.e., $SU_{gk}^D = SD_g^D \leq 1$), while slow-start technologies $\mathcal{G}^S$ are those with a SU/SD duration greater than one hour as well as a SU/SD capacity equal to the minimum power output (i.e., $SU_g = SD_g = \underline{P}_g$). Therefore, the total power output of slow-start technologies considers SU/SD trajectories (6-34), whereas (6-35) for quick-start technologies does not. For a better understanding of the modeling of quick- and slow-start technologies, the reader is referred to [71], [144].

Table 6-6. Total thermal generation output

Model	Constraint
Traditional Energy-based	$\hat{e}_{\omega gt} = \underline{P}_g u_{\omega gt} + e_{\omega gt}$ (6-32)
Energy-based using SU/SD trajectories	$\hat{e}_{\omega gt} = \underbrace{\sum_{k=1}^{K_g}\sum_{i=1}^{SU_{gk}^D} E_{gki}^{SU}\delta_{\omega gk,\left(t-i+SU_{gk}^D+1\right)}}_{Startup\ trajectory} + \underbrace{\sum_{i=1}^{SD_g^D} E_{gi}^{SD} z_{\omega g,(t-i+1)}}_{Shutdown\ trajectory} + \underbrace{\underline{P}_g u_{\omega gt} + e_{\omega gt}}_{Output\ when\ being\ up}$ (6-33)
Power-based	$\hat{p}_{\omega gt} = \underline{P}_g\left(u_{\omega gt} + y_{\omega g,t+1}\right) + p_{\omega gt} \quad \forall g \in \mathcal{G}^F$ (6-34)
	$\hat{p}_{\omega gt} = \underbrace{\sum_{k=1}^{K_g}\sum_{i=1}^{SU_{gk}^D} P_{gki}^{SU}\delta_{\omega gk,\left(t-i+SU_{gk}^D+2\right)}}_{Startup\ trajectory} + \underbrace{\sum_{i=2}^{SD_g^D+1} P_{gi}^{SD} z_{\omega g,(t-i+2)}}_{Shutdown\ trajectory} + \underbrace{\underline{P}_g\left(u_{\omega gt} + y_{\omega g,t+1}\right) + p_{\omega gt}}_{Output\ when\ being\ up} \quad \forall g \in \mathcal{G}^S$ (6-35)

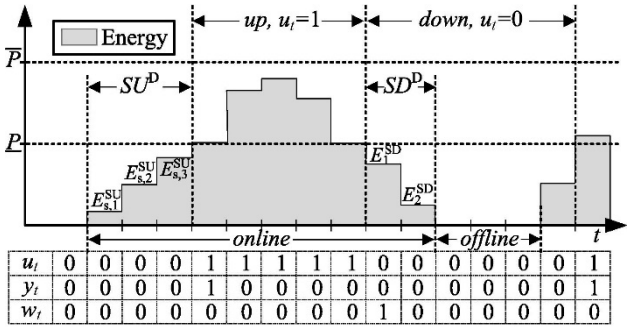


Figure 6-4. Energy-based total output with SU/SD trajectories [31]

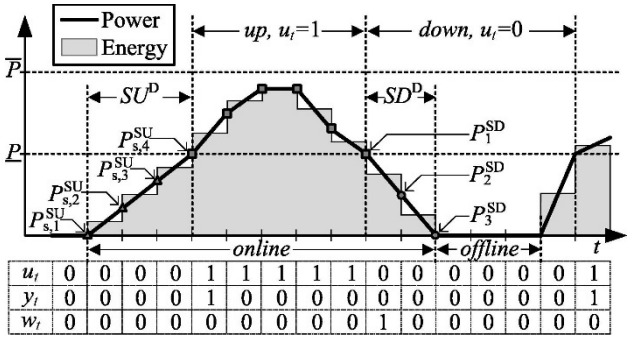


Figure 6-5. Power-based total output of thermal units [31]

The energy-based UC formulation presented here is based on the tight and compact formulation proposed in [145]. Furthermore, Gentile et al. [144] have proven that the set of constraints (6-21)-(6-23) together with (6-26),(6-28)–(6-30),(6-32) is the tightest representation (i.e., convex hull) for the energy-based model.

The power-based UC formulation presented here is based on the tight and compact formulation proposed in [70]. Furthermore, Morales-España et al.

[71] has proven that the set of constraints (6-21)-(6-23) together with (6-27),(6-31),(6-34)-(6-35) is the tightest possible representation (i.e., convex hull) for the power-based model.

6.2.7 ESS Constraints

The storage inventory level is defined with (6-36), where $\phi_{\omega st}$ is the storage level at time t for storage technology s in scenario ω and η_s is the round-trip efficiency. Constraint (6-36) is associated to an energy concept in ESS. Therefore, this constraint is the same in both models. The energy-based model considers charged ($\hat{c}_{\omega st}$) and discharged ($\hat{e}_{\omega st}$) energy as the direct output of ESS. However, in the power-based model, charged and discharged energy are determined with (6-2) and (6-3).

Constraints (6-37) and (6-38) defines the maximum and minimum storage level. EPR_s is the energy-to-power ratio of each ESS technology. For example, if we consider a BESS with a maximum energy capacity equal to 400 MWh and maximum power output equal to 100 MW, then the EPR is equal to 4h.

Notice that (6-37) and (6-38) include reserve decision variables in their definitions. In both constraints, the reserve term guarantees that enough energy will be stored in order to provide the assigned upward/downward reserve in time step t and $t-1$.

$$\phi_{\omega st} = \phi_{\omega s,t-1} + \eta_s \hat{c}_{\omega st} - \hat{e}_{\omega st} \tag{6-36}$$

$$\phi_{\omega st} \leq EPR_s(X_s^0 + x_s) - \sum_{i=t-1}^{t} r_{\omega gi}^{-} \tag{6-37}$$

$$\phi_{\omega st} \geq \sum_{i=t-1}^{t} r_{\omega gi}^{+} \tag{6-38}$$

6.2.8 Constraints for Flexibility Requirements

We focus on ramping constraints including reserve decisions for flexibility requirements in power systems. This section analyzes these constraints in both energy- and power-based model.

For the energy-based model, we analyze ramping constraints for thermal generation technologies. Figure 6-6 shows the increase/decrease on scheduled energy from time $t-1$ ($e_{\omega g,t-1}$) to t ($e_{\omega gt}$), including the upward/downward reserve ($r_{\omega gt}^{+}$, $r_{\omega gt}^{-}$). Notice that the change in the scheduled energy implies an infinite ramping capability of the thermal unit. This situation has been previously pointed out in [65], and it is one of the main reasons to develop power-based models. The power-based model overcomes this situation by explicitly defining the power trajectories instead of the hourly scheduled energies. We explain ramping constraints in the power-based model later in this section.

Instead of considering an infinite ramping capability in the energy-based model, we consider the ramping capability at τ-min (e.g., τ = 5 min). This assumption guarantees that scheduled reserves are feasible to provide at τ-min using the energy-based formulation. For instance, ramping capability limits imposed with (6-39)-(6-40) consider the reserve that thermal technologies can provide at τ-min, see Figure 6-6. Where RU_g and RD_g are the ramp-up/down capabilities in [MW/min]. When ramping constraints with τ<60 are

considered, hourly ramping constraints will be non-binding constraints. Therefore, we omit hourly ramping constraints in the energy-based model.

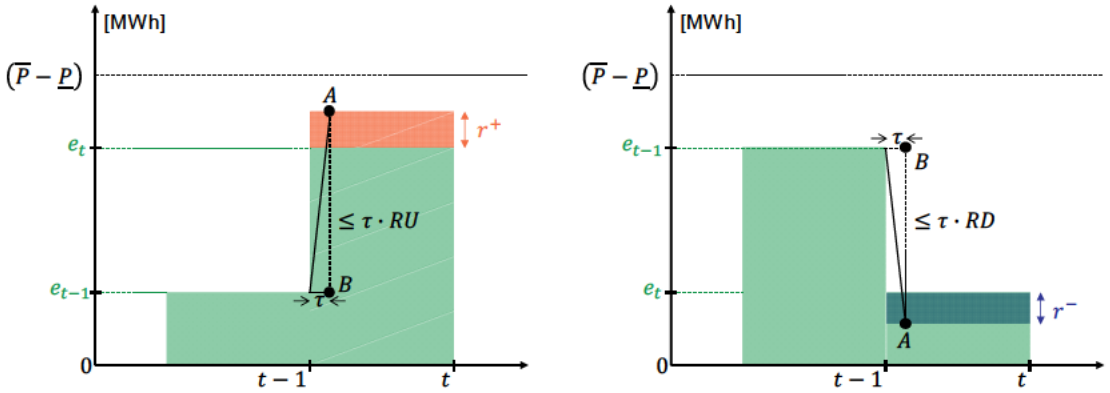


Figure 6-6. Ramping constraints in the energy-based model

$$\left(e_{\omega gt} - e_{\omega g,t-1}\right) + r^{+}_{\omega gt} \leq \tau RU_g u_{\omega gt} \tag{6-39}$$

$$\left(e_{\omega gt} - e_{\omega g,t-1}\right) - r^{-}_{\omega gt} \geq -\tau RD_g u_{\omega g,t-1} \tag{6-40}$$

For ESS ramping constraints in the energy-based model, we have a similar situation. Nevertheless, we need to consider that ESS has also a charging mode. Ramping constraints must be defined for this situation too. Figure 6-7 shows all feasible operating points for ESS. Bear in mind that we have included (6-17) and (6-19) to avoid charging and discharging at the same time. Thus, (6-41) and (6-42) guarantee the ramping constraints for ESS in all the operating points. Here, RU_s and RD_s are the ramp-up/down capabilities in [(MW/min)/MW] for ESS.

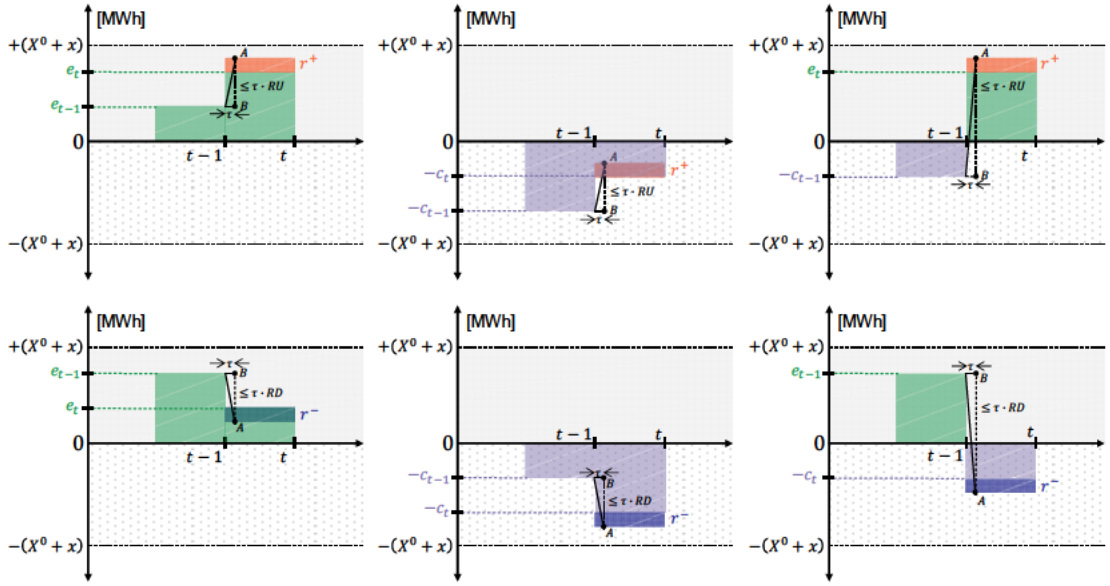


Figure 6-7. Ramping constraints for ESS in the energy-based model

$$\left(\hat{e}_{\omega st} - \hat{e}_{\omega s,t-1}\right) - \left(\hat{c}_{\omega st} - \hat{c}_{\omega s,t-1}\right) + r^{+}_{\omega st} \leq \tau RU_s (X^0_s + x_s) \tag{6-41}$$

$$\left(\hat{e}_{\omega st} - \hat{e}_{\omega s,t-1}\right) - \left(\hat{c}_{\omega st} - \hat{c}_{\omega s,t-1}\right) - r^{-}_{\omega st} \geq -\tau RD_s (X^0_s + x_s) \tag{6-42}$$

As we mentioned before in this section, one of the main advantages of the power-based formulation is that it allows describing a more detailed set of constraints to represent the flexibility requirements, which are described in

terms of power instead of energy. The proposed power-based equations in [70] ensure that reserves can be provided at any time within the hour by guaranteeing that the reserve does not exceed: 1) the ramp-capability limits at τ-min (e.g., τ=5 min), and 2) power-capacity limits at the end of the hour (i.e., 60 min) and at τ-min. Figure 6-8 shows the power trajectories between hour $t-1$ and t, including the operating points at τ-min that are guaranteed by the previous conditions. First, we need to guarantee that segments EA and AF are correspondingly below the ramp-capability limits τRU_g and τRD_g. Equations (6-43)-(6-44) guarantees this situation considering that reserves can be provided at any time within the hour. Second, we need to guarantee that points E and F are within the maximum and minimum power-capacity limit, see (6-45) and (6-46).

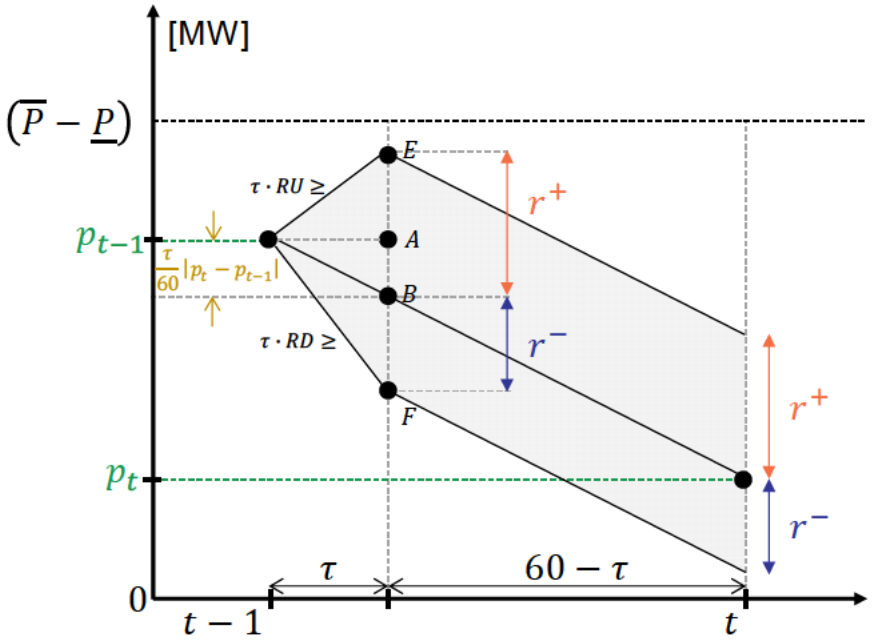


Figure 6-8. Ramping constraints in the power-based model

$$\underbrace{\underbrace{\frac{\tau(p_{\omega gt}-p_{\omega g,t-1})}{60}}_{BA}+\underbrace{r^+_{\omega gt}}_{EB}}_{EA} \leq \tau RU_g u_{\omega gt} \tag{6-43}$$

$$\underbrace{\underbrace{-\frac{\tau(p_{\omega gt}-p_{\omega g,t-1})}{60}}_{AB}+\underbrace{r^-_{\omega gt}}_{BF}}_{AF} \leq \tau RD_g u_{\omega g,t-1} \tag{6-44}$$

$$\underbrace{\underbrace{\frac{\tau p_{\omega gt}+(60-\tau)p_{\omega g,t-1}}{60}}_{B}+\underbrace{r^+_{\omega gt}}_{EB}}_{E} \leq (\overline{P}_g-\underline{P}_g)u_{\omega gt} \tag{6-45}$$

$$\underbrace{\underbrace{\frac{\tau p_{\omega gt}+(60-\tau)p_{\omega g,t-1}}{60}}_{B}-\underbrace{r^-_{\omega gt}}_{BF}}_{F} \geq 0 \tag{6-46}$$

Constraints (6-43)-(6-46) have been defined for thermal generation units in [70]. However, ramping constraints in power-based models have not been de-

fined for ESS before in the literature. This thesis tackles this issue and proposes a set of constraints for flexibility requirements in power-based models. Figure 6-9 shows different operating points and reserves for ESS at τ-min within hour t. As in the thermal generation constraints, here segments EA and AF must be below the ramp-capability limits τRU_s and τRD_s, as well as points E and F must be within the maximum and minimum power-capacity limit too. Therefore, constraints (6-47)-(6-50) guarantee these conditions for all operating points of ESS in Figure 6-9. Notice, that here it is also important to remember that we have included (6-18) and (6-20) to avoid simultaneous charging and discharging in the power-based model.

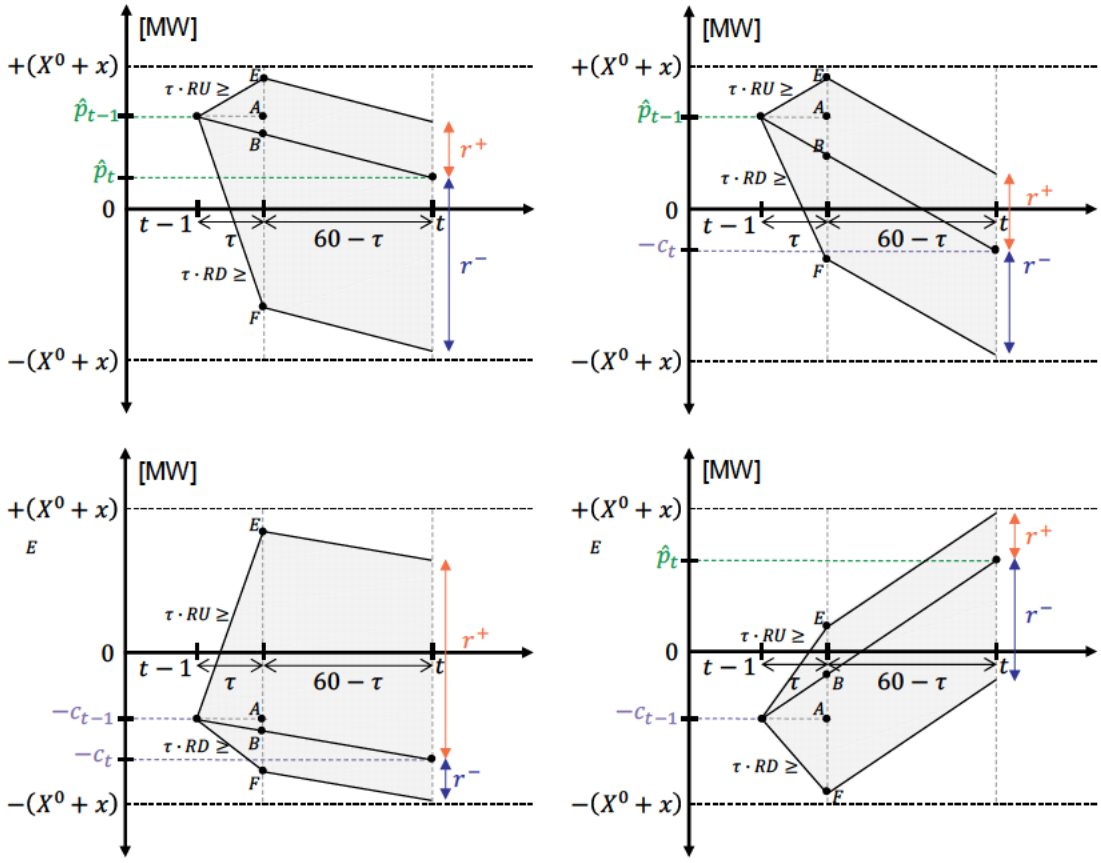


Figure 6-9. Ramping constraints for ESS in the power-based model

$$\underbrace{\underbrace{\frac{\tau(\hat{p}_{\omega st}-\hat{p}_{\omega s,t-1})}{60}-\frac{\tau(c_{\omega st}-c_{\omega s,t-1})}{60}}_{BA}}_{EA}+\underbrace{r^+_{\omega st}}_{EB}\leq \tau RU_s(X_s^0+x_s) \tag{6-47}$$

$$\underbrace{\underbrace{\frac{\tau(c_{\omega st}-c_{\omega s,t-1})}{60}-\frac{\tau(\hat{p}_{\omega st}-\hat{p}_{\omega s,t-1})}{60}}_{AB}}_{AF}+\underbrace{r^-_{\omega st}}_{BF}\leq \tau RD_s(X_s^0+x_s) \tag{6-48}$$

$$\underbrace{\underbrace{\frac{\tau(\hat{p}_{\omega st}-c_{\omega st})+(60-\tau)(\hat{p}_{\omega s,t-1}-c_{\omega s,t-1})}{60}}_{B}+\underbrace{r^+_{\omega st}}_{EB}}_{E}\leq X_s^0+x_s \tag{6-49}$$

$$\underbrace{\underbrace{\frac{\tau(\hat{p}_{\omega st} - c_{\omega st}) + (60 - \tau)\left(\hat{p}_{\omega s,t-1} - c_{\omega s,t-1}\right)}{60}}_{B} - \underbrace{r^{-}_{\omega st}}_{BF}}_{F} \geq -(X_s^0 + x_s) \tag{6-50}$$

6.3 System Flexibility Evaluation

Time resolution is an important aspect to determine the flexibility requirements in power systems. High resolution, e.g., minutes, is needed to model correctly the real operation of power systems. Nevertheless, existing GEP models use an hourly resolution under the assumption that, for long-term planning, it is enough to capture the flexibility requirements of power systems. In addition, long-term models, such as GEP, are computationally intensive problems even in an hourly resolution, therefore, if we consider lower time resolutions, it will make these problems even more difficult to solve.

Instead of reducing the time resolution, the proposed power-based GEP-UC model tackles this situation by correctly modeling the flexibility constraints and variables in Section 6.2. In order to measure the quality of the obtained solution under real-time flexibility requirements, we carry out an evaluation of investment and operational decisions through a simulation using the same scenarios as in the GEP-UC hourly optimization. This evaluation allows us to establish the problems associated to each formulation rather than those associated to the uncertainty representation by itself. This type of real-time validation is not commonly carried out because it is considered unnecessary. Nevertheless, to validate correctly flexibility capabilities and requirements of the system, this real-time evaluation is crucial [31].

The complete procedure to calculate investment decisions and ex-post real-time evaluation is shown in Figure 6-10 (top). During stage 1, the investment and hourly UC schedule are optimized solving the formulations shown in Section 6.2. Then, investment, commitment, and reserve decisions are fixed. Stage 2 tests the results through a real-time simulation model, using a 5-min optimal dispatch (emulating real-time markets as in [31]) in order to evaluate the GEP-UC solution. Dispatch decisions (e.g., production, charge/discharge) obtained in stage 2 are called redispatches, allowing us to evaluate the deviations with respect to the stage 1. This is referred to as the integer approach. In addition, we propose a semi-relaxed approach for the power-based formulation, which is shown in Figure 6-10 (bottom). Here we split stage 1 in two sub-stages. First, the sub-stage 1a solves the power-based formulation considering integer investment decisions and continuous decisions for UC ($u_{\omega gt}$, $y_{\omega gt}, z_{\omega gt}$) variables and ESS ($\gamma_{\omega st}$) charging/discharging variable. This approximation allows solving the GEP problem much faster. Then investment decisions are fixed in sub-stage 1b, where the power-based formulation is solved considering integer UC decisions. Once again, investment, unit commitment, and reserve decisions are fixed to simulate a 5-min optimal dispatch.

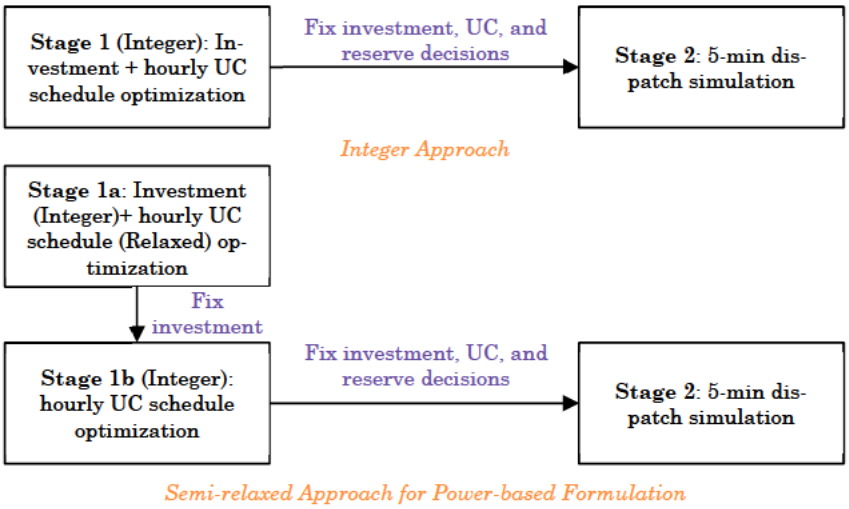


Figure 6-10. Stage sequence for both approaches

6.4 Case Studies and Optimization Models

To evaluate the performance of the different approaches, we use two case studies: a modified IEEE 118-bus test system and a stylized Dutch power system in target year 2040. Input data for both case studies is available online at [146], including the 5-min demand and renewable production profiles. Both case studies are solved considering a green-field investment approach (i.e., no initial capacity) for thermal generation and ESS investment, while the VRES capacity is predefined.

The modified IEEE 118-bus test system is described in Morales-España [65] for a time span of 24 h. This system was originally conceived for UC problems and it has 118 buses, 186 transmission lines, 91 loads, 54 slow-start thermal technologies, 10 quick-start technologies, and three buses with wind production. Nevertheless, we adapt this case study for GEP problems. Thermal unit investments are allowed in buses where there was a unit connected in the initial UC problem. In addition, ESS investment decisions are available in three types of technologies (PSH, CAES, and BESS) for buses with renewable production. The total (5-min) load average is 3578.6 MW, it has a peak of 5117.5 MW and a minimum of 1435.4 MW.

The stylized Dutch system case study for year 2040 is mainly based on the information available in the Ten Year Network Development Plan 2018 [147] (e.g., hourly demand profile, renewable capacity, technical characteristics and available technologies). However, the wind and solar profiles were taken from [101], [102] since this information is not available in [147]. Instead of solving 8760 h for the whole year, we have selected four representative weeks using the proposed method in [95] and k-medoids clustering technique [148]. Other authors [57], [149] have proposed different approaches to select the representative periods (e.g., weeks or days) that are compatible with the proposed GEP-UC models in this paper. Each representative week is considered as one scenario in the optimization problem, and the scenario probability is obtained from the clustering process. For investment decisions, four different thermal generation technologies are considered, Combined Heat and Power (CHP), combined cycle gas turbine (CCGT), open cycle gas turbine (OCGT), and Light

Oil (Oil). Moreover, three ESS (PSH, CAES, BESS) technologies are considered for investment decisions.

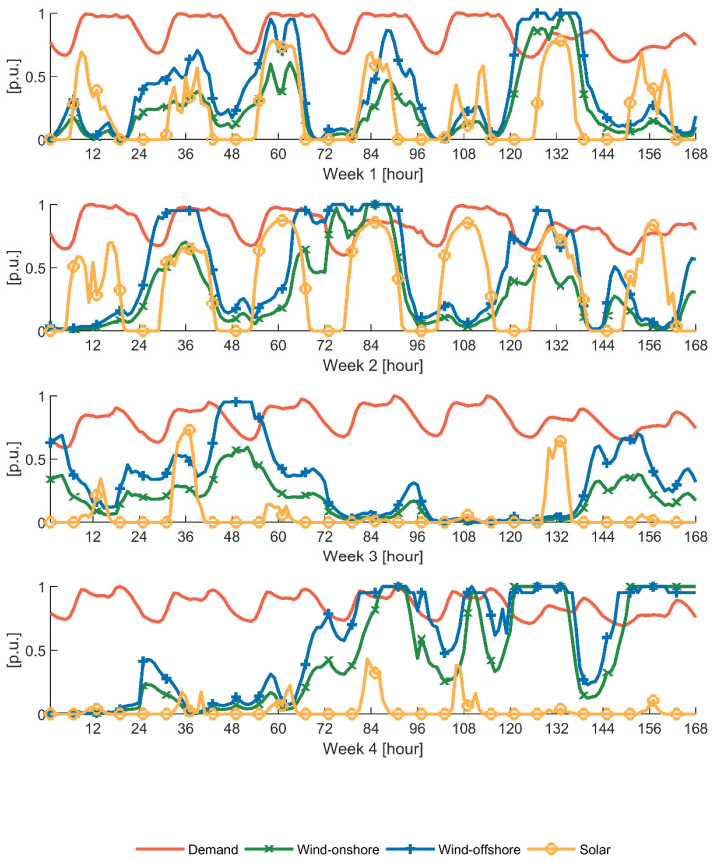


Figure 6-11. Representative weeks for Dutch case study

For each case study, four different models are implemented: traditional energy-based (EB), energy-based including SU/SD power trajectories (EBs), the proposed power-based formulation (PB), and the semi-relaxed power-based formulation (SR-PB). Table 6-7 shows the summary with all the implemented models. All models consider τ = 5min for constraints associated to flexibility constraints.

All optimizations were carried out using Gurobi 8.1 on an Intel®-Core™ i7-4770 (64-bit) 3.4-GHz personal computer with 16GB of RAM memory. The problems are solved until they reach an optimality tolerance of 0.1%.

Table 6-7. Summary of GEP-UC models

Equations	EB	EBs	PB	SR-PB
Objective function	(6-1)			
System constraints	(6-4),(6-6),(6-8),(6-9)		(6-5),(6-7)–(6-9)	
Investment constraints	(6-10),(6-11),(6-13),(6-15)		(6-10),(6-12),(6-14),(6-16)	
Unit Commitment constraints	(6-21)–(6-25)			

Equations	EB	EBs	PB	SR-PB
Thermal generation technology constraints	(6-26),(6-28)–(6-30)		(6-27),(6-31)	
Total output thermal technologies	(6-32)	(6-33)	(6-34),(6-35)	
ESS constraints	(6-36)–(6-38)			
Constraints for flexibility requirements in thermal technologies	(6-39),(6-40)		(6-43)–(6-46)	
Constraints for flexibility requirements in ESS	(6-17),(6-19),(6-41),(6-42)		(6-18),(6-20),(6-47)–(6-50)	
Integer variables	$u_{\omega gt}$, $y_{\omega gt}$, $z_{\omega gt}$, $\gamma_{\omega st}$, $\delta_{\omega gkt}$, x_j			Stage 1a: x_j Stage 1b: $u_{\omega gt}$, $y_{\omega gt}$, $z_{\omega gt}$, $\gamma_{\omega st}$, $\delta_{\omega gkt}$

6.5 Results

6.5.1 Modified IEEE 118-bus Test System

Table 6-8 shows the main results for each model. The total investment cost (ESS + Thermal) is higher in the classic EB model than the one obtained with the PB model. Generally, increasing investments lowers operating cost. Nevertheless, here we obtain a counterintuitive result. Even though the classic EB model invests more (6%), the operating cost is higher than the one in the PB model (15%). Moreover, the CO2 emissions and curtailment are also higher in the classic EB model, despite its higher capacity in clean ESS and lower capacity in thermal technologies. This is also a counterintuitive result, because at first glance, less thermal generation should pollute less, and more storage should allocate more renewables. However, this result is related to how the technology mix is selected in each model. Therefore, it is not only a matter of how much the model invests, it is also a matter of how the technology mix is selected, see Table 6-9. For instance, although the total coal capacity is higher in the proposed PB model, the actual total coal production is lower (7%) than the one in the classic EB model, see Table 6-10. This is compensated by a higher use of wind, gas (that have a lower CO2 emission factor) and oil, which overall results in lower CO2 emissions. As mentioned in Section 6.1 the PB model equations allow scheduling the thermal technologies in a way that correctly represents the requirements and actual availability of the system's flexibility, such as the load ramps. The results show the benefits of accurately considering the flexibility requirements and of correctly modeling the flexibility capabilities of the system by modeling in terms of power instead of energy.

The EBs model improves the classic EB model by including the SU/SD power-based ramps. In stage 1, the total cost in the EBs model is 8.5% lower than the classic EB model. However, it is still 4% higher than the PB model and with more curtailment (5.7 times). The EBs technology mix is also different, as it invests more in PHS and coal (Table 6-9). And yet, the PB model allocates more wind with less ESS, see Table 6-10. Therefore, the PB model invests more efficiently due to a more accurate representation of flexibility requirements and capabilities of the power system.

Regarding the CPU time, the PB model is faster than its energy counterparts (2.4 and 1.5 times respectively). Nevertheless, for large-scale investment decision problems, the integer nature of the UC variables especially could make the problem intractable to solve. Therefore, the proposed SR-PB models aims at overcoming this difficulty. For instance, it solves the problem 9 times faster than the PB model and with only a 0.2% difference in the objective function. Moreover, the difference in the CO2 emissions is only 0.4%. The main difference appears in the curtailment (90%) due to the increase in the investment made by the SR-PB that allows reducing the operating cost by increasing wind production. When the SR-PB and the EB are compared, it may be concluded that even the semi-relaxed version of the power-based model (i.e., SR-PB) shows better performance than the discrete version of the energy-based models (i.e., EB and EBs). In other words, the SR-PB model has a lower total cost than the EB model, investing and operating with lower cost, while simultaneously solving 21 times faster. Although this result was obtained for the IEEE-bus test system, we have observed in other case studies (i.e., the Dutch power system in Section 6.5.2) that this result seems to hold.

Table 6-8. IEEE 118-bus System: Performance for each formulation

	Result	EB	EBs	PB	SR-PB
Stage 1	Total Cost [M$]	10.15	9.29	8.94	8.96†
	ESS Invest Cost [M$]	0.43	0.35	0.19	0.17
	Therm. Invest Cost [M$]	1.01	1.42	1.17	1.24
	Operating Cost [M$]	8.71	7.52	7.58	7.55†
	CO2 emissions [ton]	63.11	53.06	53.98	53.74
	Curtailment [%]	5.76	4.18	0.73	0.70
	CPU Time [s]	10717	6767	4478	500
Stage 2	Operating Cost [M$]	8.22	7.53	7.58	7.55
	Total Cost [M$]	9.66	9.30	8.94	8.96
	CO2 emissions [ton]	59.31	52.48	53.95	53.71
	Curtailment [%]	0.00	0.00	0.60	0.62

† Values from Stage 1b

The results in Table 6-8 for stage 2 are also showing interesting information: comparing the operating cost between stage 1 and 2, the classic EB shows a reduction of 6%, while in the other models remain almost the same. Moreover, the curtailment is also reduced from stage 1 to stage 2 in both energy-based models, while it remains almost the same in the power-based models. These reductions in the results suggest that the obtained schedule in stage 1 with energy-based models leads to more re-dispatches in the technologies in stage 2. Figure 6-12 illustrates this situation with the deviation with respect to the hourly thermal production obtained in stage 2 for each model. In both energy-based models, downward deviations are higher than upward deviations, which explains why the operating cost is reduced from stage 1 to stage 2 in the classic EB model as well as the reduction on the curtailment for both energy-based models. The power-based models show deviations in both directions lower than 3%, which means that the hourly schedule (stage 1) is better fitted for the 5-min real-time operation (stage 2). This high deviation of the energy-based models is due to their intrinsic incapability to accurately represent the flexibility needs and capabilities. These conclusions are aligned with

those in [31] where different case studies where carried out disregarding investment decisions.

Table 6-9. Technology investment decisions [MW]

Technology	EB	EBs	PB	SR-PB
PSH	1250	1000	500	441
CAES	0	0	0	0
BESS	150	150	150	150
GAS	360	600	420	480
COAL	4380	6080	5030	5330
OIL	50	100	100	100

Table 6-10. Technology production decisions [MWh]

Technology	EB	EBs	PB	SR-PB
PSH	7352	5944	2449	2019
CAES	0	0	0	0
BESS	1053	1003	1035	1033
GAS	494	2719	2482	2680
COAL	67540	63939	62913	62570
OIL	52	900	950	900
WIND	18880	19196	19887	20018

Notice that ESS plays an important role in the reschedules made in stage 2. Therefore, we run a sensitivity case in which the State-of-Charge (SoC) at the end of each hour is a lower bound for the ESS in the stage 2. This limits the reschedules made in this stage, increasing the operating cost. Table 6-11 shows that situation, where with this additional constraint the operating cost, CO2 emissions and curtailment are higher than in the base case. It is important to highlight that in this sensitivity case energy-type models cannot reduce the curtailment to zero like in the base case. Therefore, the flexibility provided by the ESS was partly responsible for the reduction of the curtailment between stage 1 and 2 in this type of models. Figure 6-13 shows the SoC in the batteries during stage 2 for the base case and the sensitivity case. The difference between both results in each model shows how the energy-type models were taking advantage of the ESS to reduce the operating cost in stage 2 at the cost of more rescheduling in the thermal technologies.

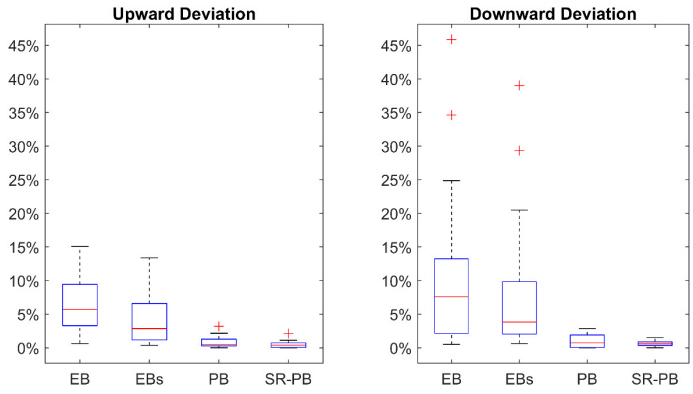


Figure 6-12. Stage 2 deviation in scheduled thermal output

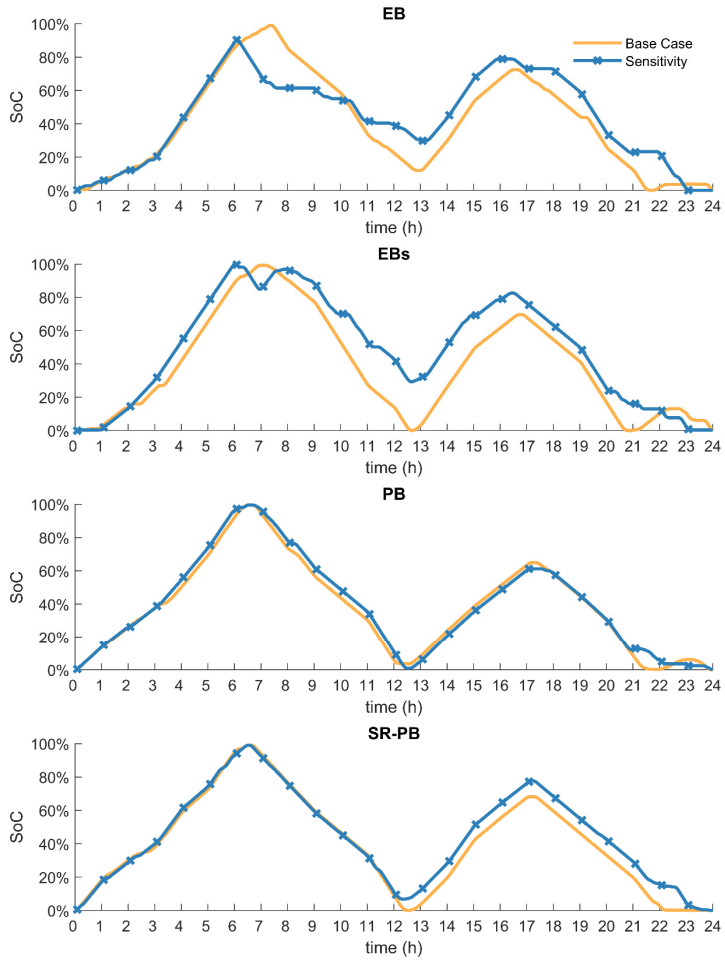


Figure 6-13. BESS SoC in Stage 2 obtained for each model

Table 6-11. IEEE 118-bus System: Stage 2 – sensitivity results

	Result	EB	EBs	PB	SR-PB
Stage 2	Operating Cost [M$]	8.35	7.66	7.60	7.55
	Total Cost [M$]	9.79	9.43	8.96	8.96
	CO2 emissions [ton]	59.89	52.71	54.04	53.73
	Curtailment [%]	1.99	0.98	0.62	0.13

6.5.2 Dutch Case Study without VRES Investment

Table 6-12 shows the results for a stylized Dutch power system. The main conclusions drawn from the previous case study remain valid. That is, the classic EB model obtains the most expensive investment, and the operating cost is also the highest, while also resulting in the highest CO2 emissions. The amount of ESS invested in the EB model is also the highest, hence allowing it to obtain less curtailment than PB in the stage 2. Nevertheless, still the PB model results in the lowest total cost in both stages and solves the GEP problem faster than EB. Therefore, modeling flexibility requirements for the Dutch case study with the PB model also leads to a better solution than the classic EB model. In addition, the SR-PB further reduces the CPU time without losing accuracy in the results (less than 0.1% in total cost).

Table 6-12. Stylized Dutch System: Performance for each formulation

	Result	EB	EBs	PB	SR-PB
Stage 1	Total Cost [M$]	73.18	70.39	68.14	68.16†
	ESS Invest Cost [M$]	13.47	11.15	10.53	10.88
	Therm. Invest Cost [M$]	13.79	14.12	13.43	13.47
	Operating Cost [M$]	45.92	45.12	44.18	43.81†
	CO2 emissions [ton]	112.10	98.06	89.46	88.77
	Curtailment [%]	44.72	45.47	45.46	45.34
	CPU Time [s]	571	161	131	60
Stage 2	Operating Cost [M$]	45.76	46.61	44.90	44.52
	Total Cost [M$]	73.02	71.88	68.86	68.87
	CO2 emissions [kton]	107.73	100.01	94.29	93.44
	Curtailment [%]	47.88	48.35	48.34	45.39

† Values from Stage 1b

Figure 6-14 shows the optimal technology mix for each model. The EB model invests 16% more in ESS capacity than the PB model, however, this is not leading to a more flexible system (less operating cost, see Table 6-12) because it also invests in less flexible thermal technologies such as CHP. However, the CHP investment disappears if the SU/SD trajectories are considered. This leads to more investment in CCGTs in the EBs model, making the technology mix more similar to the one obtained in the PB model. The main difference remains in the ESS invested technologies; the power-based models invest in both PSH and BESS, while both energy-based models only invest in BESS technology. This is mainly due to the overestimation of the flexibility made by the energy-based models, which leads to a more expensive operation in the real time validation (i.e., stage 2 in Table 6-12).

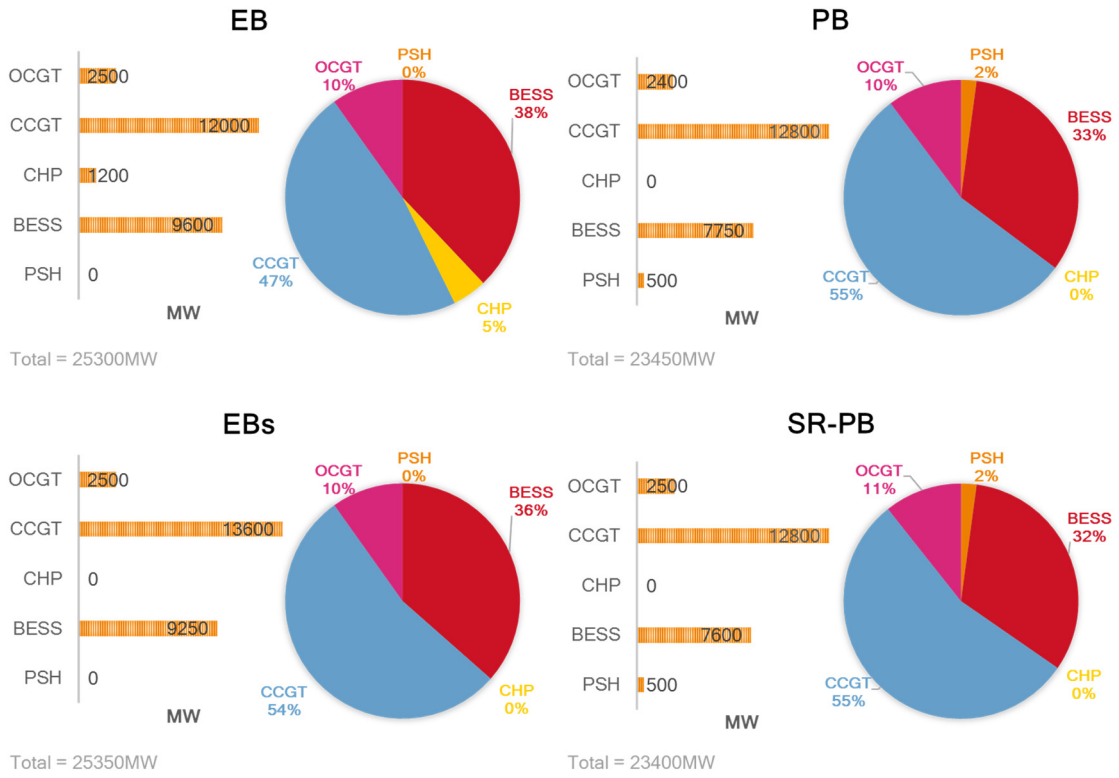


Figure 6-14. Generation mix - Dutch case study

The approximation made by the SR-PB model is also close to the obtained technology mix in the PB model. It also obtains a combination of BESS and PSH for the ESS, as well as a combination of OCGT and CCGT for the thermal generation technologies. Therefore, the SR-PB model approximates quite well not only the objective function (i.e., total cost), but also the technology capacity mix.

6.5.3 Dutch Case Study Sensitivity to Ramp Capacity

In addition to the base case shown in the previous section, Table 6-13 shows a sensitivity where ramp capabilities of thermal technologies are twice than before, i.e., thermal technologies are now much more flexible. As the flexibility of the thermal resources increases, the difference between energy-based and power-based models decreases. For instance, the difference between the EB the PB models changes from 7.4% to 4.3%. Therefore, if the power system does not have ramping problems, i.e., flexibility is not a problem in general, the difference between energy-based and power-based models is less significant. However, if flexibility is a limited resource and needs to be correctly managed, then the power-based models are the right option to obtain the capacity expansion planning for the system.

Table 6-13. Stylized Dutch System: Sensitivity to Ramp Capacity

	Result	EB	EBs	PB	SR-PB
Stage 1	Total Cost [M$]	70.51	67.93	67.60	67.61†
	ESS Invest Cost [M$]	13.35	10.97	10.66	10.74
	Therm. Invest Cost [M$]	13.47	13.47	13.43	13.47
	Operating Cost [M$]	43.69	43.49	43.51	43.40†
	CO2 emissions [ton]	103.10	90.84	88.46	88.17
	Curtailment [%]	44.32	45.22	45.16	45.19
	CPU Time [s]	142	130	100	43
Stage 2	Operating Cost [M$]	44.37	45.92	44.25	44.21
	Total Cost [M$]	71.19	70.36	68.34	68.42
	CO2 emissions [kton]	100.76	94.35	93.07	93.07
	Curtailment [%]	44.62	45.30	45.24	45.28

† Values from Stage 1b

6.5.4 Dutch Case Study Analysis of VRES Curtailment Cost

The base case considers a predefined VRES capacity and zero penalization to VRES curtailment. Some countries have implemented the so-called feed-in tariff in order to incentive VRES investments [150]. One consequence of feed-in tariffs is that VRES receive a payment no matter if they produce energy or not. In other words, the power system also pays for the curtailment. This is represented through a penalization in the objective function of the models in this section. Here we analyze the impact on the generation technology mix with different curtailment penalizations. Figure 6-15 shows the ESS share for all models, including the value in MW for each technology. In all models, the generation technology mix changes as the penalization increases too. In addition, the higher the penalization the more ESS capacity is installed. This is mainly because it is cheaper to invest in more ESS than paying the curtailment penalization cost. Therefore, forcing VRES production through curtailment penalizations leverages up the ESS investment. The PB model results

show that the higher the curtailment penalization the higher the investment in CAES, which is the technology with the least efficiency coefficient. Therefore, forcing renewable production could lead to an ESS mix in which the least efficient (but cheaper) technologies are favored.

The results in this section show an important policy insight for regulatory authorities in order to develop efficient incentive schemes that may promote ESS investment.

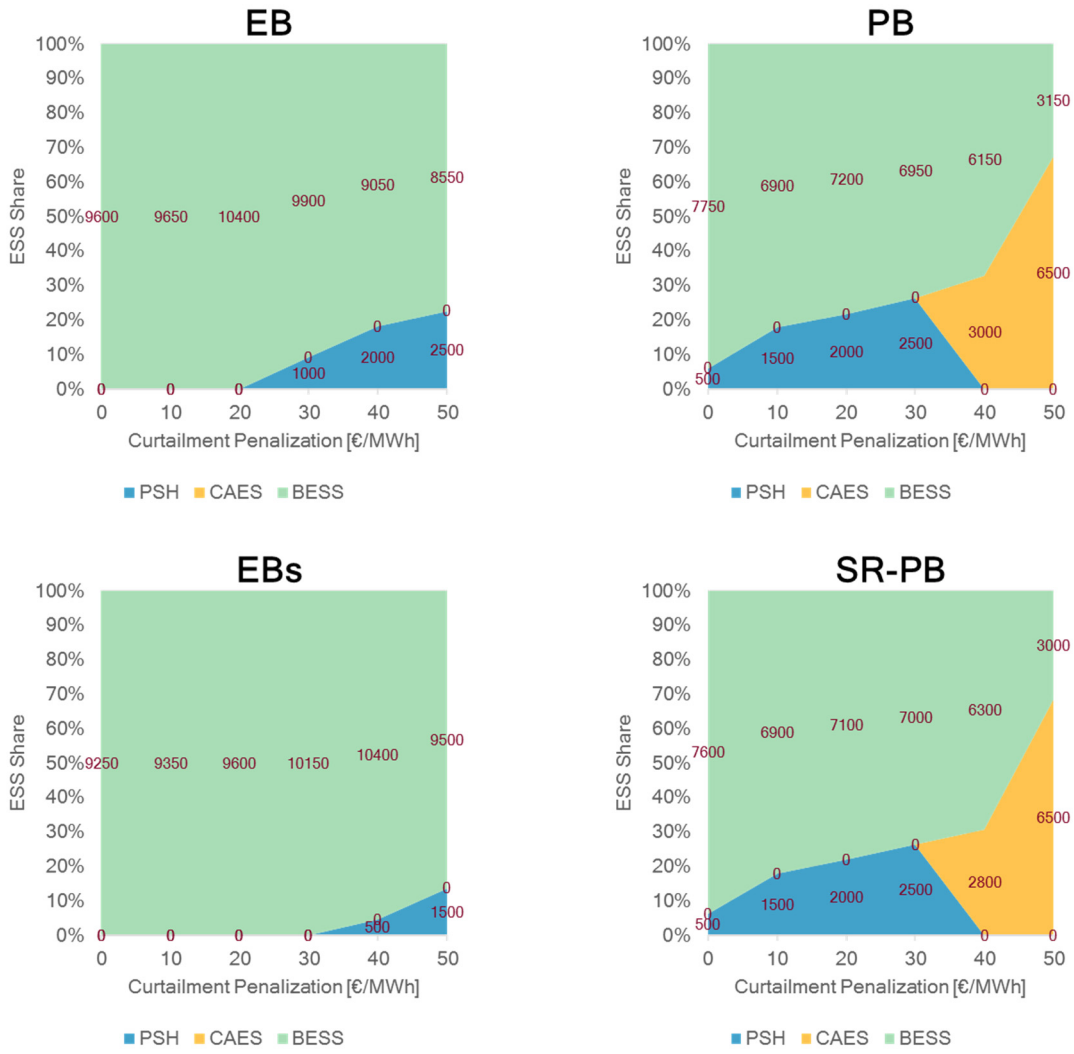


Figure 6-15. Sensitivity to curtailment cost

6.5.5 Dutch Case Study with VRES Investment

The previous case study had a predefined VRES capacity, and VRES investments were not allowed. In this section, we allow VRES investment in order to validate if previous conclusions are still consistent. In this section, no ESS or VRES limits were imposed on their investment decisions.

Table 6-14 shows the results for all models. Here, the total investment cost (i.e., ESS, VRES, and thermal generation) is 4.5% higher in the EB model than the one obtained in the PB model. Despite the EB model invests more, the operating cost is 2.9% higher compared to the PB model. Moreover, the EB operates with more CO2 emissions (12.6%). Therefore, previous conclusions are still valid, that is, the PB model invests less, operates at lower cost,

and its operational decisions are more environmentally friendly. Even in its relaxed version, the SR-PB model has a better performance (i.e., less operating cost and less CO2 emissions) in the real-time validation (i.e., stage 2) than the EB model. In this case study, the SR-PB model also shows an excellent approximation to the PB model because the difference between the total cost in both models is less than 0.1%.

The EBs model, which includes the SU/SD trajectories, has a better performance than the EB model. It also invests less (1.2%), operates at a lower cost (2.5%), and has less CO2 emissions (11.6%). Nevertheless, it has a higher value in the total cost (2.2%) than the ones obtained in both power-based models. Despite the EBs models improves the traditional EB model, it is still worse than the power-based models.

Table 6-14. Stylized Dutch System including VRES investment

	Result	EB	EBs	PB	SR-PB
Stage 1	Total Cost [M$]	168.95	166.03	162.45	162.56†
	ESS Invest Cost [M$]	14.41	13.48	10.74	10.74
	VRES Invest Cost [M$]	65.78	65.55	66.22	65.74
	Therm. Invest Cost [M$]	18.42	18.42	17.20	17.30
	Operating Cost [M$]	70.34	68.58	68.29	68.78†
	CO2 emissions [kton]	117.45	103.86	102.60	102.24
	Curtailment [%]	22.48	22.62	23.60	23.36
	CPU Time [s]	290.06	234.83	225.48	180.88
Stage 2	Operating Cost [M$]	69.36	69.61	68.85	69.29
	Total Cost [M$]	167.97	167.06	163.01	163.07
	CO2 emissions [kton]	113.28	106.36	106.54	106.75
	Curtailment [%]	22.46	22.48	23.59	23.34

† Values from Stage 1b

Figure 6-16 shows the generation capacity mix results for all models. The total share of VRES is 4% higher in the PB model, while the investment in flexible resources (i.e., ESS) is 4% lower. And even so, the EB model invests more to operate at a higher cost in the real-time validation with more pollution (CO2 emissions), see results of stage 2 in Table 6-14. It is also important to highlight that the ESS technology investment is very different in energy-based models compared to the power-based models. For instance, the PB model invests almost 50% of ESS capacity on both BESS and PSH, while the EB model invest almost 74% of ESS capacity on BESS and 26% on PSH. Therefore, the main differences between both models are not only expressed in terms of the operating cost but also in terms of the optimal ESS capacity mix.

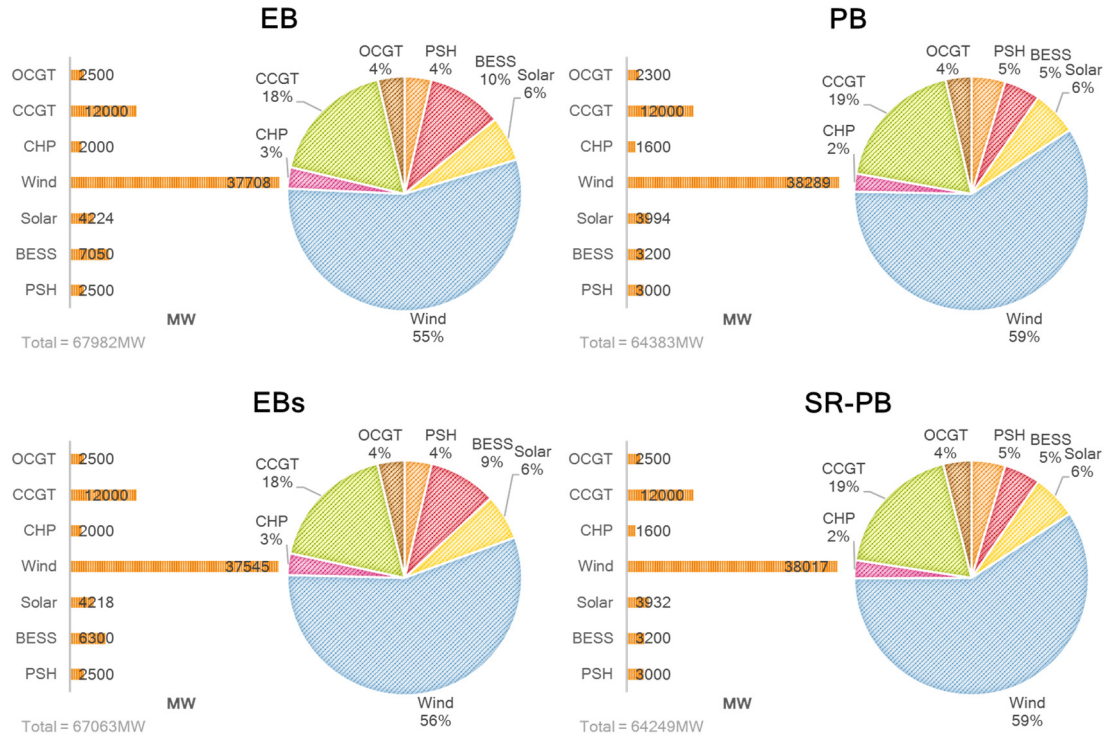


Figure 6-16. Generation mix including VRES investments

6.5.6 Dutch Case Study Sensitivity Limiting Investments

There is a close relationship between the VRES and ESS investments. Therefore, if we limit the VRES investment then the ESS investment is also limited, and vice versa. We explore this relationship in all models by running two sensitivities, one limiting the VRES investment (on-shore wind is limited to 7400 MW, off-shore wind is limited to 22466 MW, and solar PV is limited to 26455 MW)[15] and another limiting the ESS investment (BESS is limited to 5000 MW, and PSH is limited to 2500 MW).

Table 6-15 shows the different costs in the objective function for each model in the base case, and both sensitivities. In addition, Table 6-16 shows the ESS and VRES capacity for each model.

On the one hand, the EB and EBs models are more sensitive to the limitation on ESS investment because their ESS mix investment is more biased to one technology, see Table 6-16. Therefore, if that technology (in this case the BESS) is limited the ESS investment has a bigger change. On the other hand, the power-based models have an ESS technology mix in the base case more distributed between the BESS and the PSH, therefore, the impact in the ESS investment due to the limitation is lower (less than 1%), and even the total cost increases due to change in the optimal ESS technology mix. In all models, the VRES investment is reduced due to the limitation on ESS.

Limitations on VRES investment impacts more the power-based models than the energy-based models. For instance, the PB model reduces a 14.6% the ESS investment cost due to an increase of 3.8% of VRES investment, while the EB model changes 5.6% the ESS investment cost due to an increase of

[15] These values were obtained from ENTSO-E *"Ten Year Network Development Plan 2018"* [147]

6.2% in the VRES investment cost. This is mainly because the VRES technology mix is changed to more solar PV technology (66% in the PB model, see Table 6-16), which favors ESS with a daily cycle such as the BESS. This change is less significant in the EB models because in the base case there were mainly BESS investment, therefore, there were already enough BESS to deal with the daily cycles of solar PV technology. However, the PB model is still better because it invests less and obtains a lower operating cost than the EB model. This situation also occurs in the ESS investment limitation sensitivity. Therefore, the main conclusions comparing the EB and PB model are still valid for these sensitivities.

Table 6-15. Sensitivity to ESS and VRES investment - Costs

Result	Model	No limit to ESS and VRES investment	No limit to ESS and limit to VRES investment	No limit to VRES and limit to ESS investment
ESS Invest Cost [M$]	EB	14.41	13.60	11.86
	EBs	13.48	12.28	11.86
	PB	10.74	9.17	10.92
	SR-PB	10.74	8.61	10.86
VRES Invest Cost [M$]	EB	65.78	69.83	65.11
	EBs	65.55	70.49	64.42
	PB	66.22	68.72	65.61
	SR-PB	65.74	68.62	65.69
Therm. Invest Cost [M$]	EB	18.42	18.42	18.42
	EBs	18.42	18.42	18.42
	PB	17.20	17.82	17.77
	SR-PB	17.30	17.25	17.68
Operating Cost [M$]	EB	70.34	75.59	74.02
	EBs	68.58	73.37	71.52
	PB	68.29	75.20	68.23
	SR-PB	68.78†	76.41†	68.26†
Total Cost [M$]	EB	168.95	177.44	169.40
	EBs	166.02	174.56	166.22
	PB	162.45	170.91	162.54
	SR-PB	162.56†	170.90†	162.50†

† Values from Stage 1b

Table 6-16. Sensitivity to ESS and VRES investment - Capacity

Result	Model	No limit to ESS and VRES investment	No limit to ESS and limit to VRES investment	No limit to VRES and limit to ESS investment
BESS Capacity [MW]	EB	7050	7300	5000
	EBs	6300	7150	5000
	PB	3200	3750	4250
	SR-PB	3200	3500	3800
PSH Capacity [MW]	EB	2500	2000	2500
	EBs	2500	1500	2500
	PB	3000	2000	2500
	SR-PB	3000	2000	2500

Result	Model	No limit to ESS and VRES investment	No limit to ESS and limit to VRES investment	No limit to VRES and limit to ESS investment
Solar PV Capacity [MW]	EB	4224	6268	4486
	EBs	4218	6990	3861
	PB	3994	6644	3810
	SR-PB	3932	6718	3940
Wind Capacity [MW]	EB	37708	21507	36924
	EBs	37545	21339	37153
	PB	38289	20939	38066
	SR-PB	38117	20869	38027

6.6 Main Takeaways in this Chapter

This chapter proposes a power-based model to determine the GEP that improves the existing energy-based GEP models. The proposed model optimizes investment decisions on VRES, ESS, and thermal technologies. In addition, it includes real-time flexibility requirements, and the flexibility provided by ESS, as well as other UC constraints, e.g., minimum up/down times, startup and shutdown power trajectories, network constraints. The main analysis and results can be summarized in the following conclusions:

- *In the case studies that have been analyzed in this chapter, the proposed power-based model uses the installed investments more efficiently and more effectively since: 1) it represents the reality of flexibility requirements of the power system more adequately; and 2) it adequately exploits the flexibility capabilities of the system. That is, the decisions made with the power-based model simultaneously yield lower investment costs, operating cost, CO2 emissions, and renewable curtailment with respect to the energy-based model. This is mainly because the energy-based model overestimates flexibility capabilities, failing to capture the flexibility requirements such as load and VRES ramps even in a deterministic approach (i.e., without uncertainty on demand, or renewable production).*
- *The advantages of the power-based approach could become much more significant considering uncertainty. Therefore, correctly modeling the system flexibility changes the optimal expansion capacity decisions. For instance, the power-based model obtains less total investment because it is more accurate in the representation of ramping characteristics for generation resources (e.g., thermal technologies and ESS), which leads to less operating cost in the real-time validation.*
- *The power-based model has computational advantages in terms of CPU time. The results for the case study show that the power-based model is 1.3 to 4 times faster than the energy-based model. We also have shown that the semi-relaxed power-based model is even faster (10 to 21 times) without losing accuracy in the results compared with the non-relaxed power-based model (less than 0.2% objective function error). This is relevant for applications with large-scale long-term capacity expansion planning problems where relaxed models are more often used due to computational power limitations.*

- *Wrong incentives may lead to less efficient solutions. That is, forcing renewable production leads to a change in the optimal flexible resources the system needs, making lower round-trip efficiency ESS technologies more attractive, such as CAES or PSH. This means consuming more energy in the charging process to avoid penalization due to VRES curtailment. This effect is independent of the model (either energy- or power-based), however, the ESS technology mix changes radically depending on the model. For instance, the Dutch case study, considering 50€/MWh, shows that the power-based GEP-UC model invests 71% less in BESS, and invest in CAES instead of PHS, in comparison to the energy-based GEP-UC model.*
- *The results show an important insight for ISOs because, even without uncertainty, the current energy-based models impose more rescheduling in the real-time operation than the power-based models. For planning authorities this is also important because decisions made with power-based models lead to a generation technology mix that is better adapted to real-time system operation.*
- *Although the case studies that have been analyzed in this chapter have not included LTESS (e.g., hydro generation), the power-based models are compatible with the enhanced version of representative periods (i.e., RP-TM&CI in chapters 3 and 5) in order to co-optimize both types of ESS including the flexibility requirements that power-based models have.*

7 CONCLUSIONS

The relevance of this thesis is being at the cutting edge of knowledge about the use of energy storage technologies in electric power systems. Although the technology itself already exists, there are still no electrical power systems that are operated mostly with renewables and energy storage systems. It is not yet known what repercussions this type of operation would have for the electric power system, for the electricity market, economic viability, and power system operation. If we want to reach an electricity sector free of emissions, we must improve the optimization models that help in the decision-making process. Through this thesis, we have proposed several improvements to current models that may help system operators, regulatory authorities, policy makers, universities, and research centers to analyze in a more accurate way the role of energy storage systems in the electric power industry decarbonization.

7.1 Summary of Main Results

The remainder of this section contains a summary of the main results of this thesis and the most relevant conclusions that can be drawn from the presented research, according to the main chapters in this thesis.

Representation of Energy Storage Operation

- Having extended the system states to a framework with a transmission network, we have found that if the transmission network is constrained, the accuracy of the system state approach for representing storage operation improves. This conclusion is quite relevant and some-how counterintuitive, because it means that it is not necessary to include information of the congestion of the network in the clustering-process to improve the results when the system state framework is used to represent ESS operation in medium- and long-term planning models.
- The system states methodology was originally developed to include chronology and high time resolution details in operational planning models. While it can deal with long-term storage (LTESS), it cannot accurately estimate short-term storage (STESS), and quickly becomes computationally intractable because of the storage constraints. Unlike the system states models, the traditional representative periods (RP) model cannot handle long-term storage (LTESS), however, it deals well with short-term storage (STESS) as it preserves within-day chronology.
- The proposed enhanced version of the RP model (RP-TM&CI model) combines aspects of the system states and representative period models to account for both short and long-term storage. According to the case study results, it is the most accurate of the analyzed models and does not require a significant increase of CPU time.
- These results support the idea that including chronological information among representative periods may be an efficient way to include small time scale variations in longer-term planning models that involve storage. Doing so is a critical need in the adequate representation of power systems that include a significant and increasing quota of variable renewable sources and energy storage systems.

Co-optimization of ESS in hydro-thermal dispatch models

- The results validate the initial hypothesis that short-term energy storage (e.g., BESS) decisions on energy production impact the opportunity cost (or water value) of seasonal storage. This is relevant because traditional hydrothermal dispatch models disregard this situation in their formulations.
- The proposed hydrothermal model using the Linked Representative Periods (LRP) obtains a better estimation of operational results (e.g., productions, number of cycles for short-term storage, and storage levels) than the classic Load Duration Curve (LDC)models for all the case studies in the analysis. These results are expected to be in the same way for different case studies, since the LRP is a more detailed model to represent the energy storage operation than the LDC model.
- Hourly opportunity costs internalize long-term signals due to seasonality in the power system. In other words, the water value in seasonal storage includes the impact of short-term operational decisions.
- This is important to help market participants or planning authorities in their decision-making process (bids or investment decisions) by determining correct economic signals (i.e., short-term prices and long-term expected values) with the co-optimization approach in the LRP model thereby avoiding sub-optimal solutions from iterative processes (e.g., fixing the hydro reservoirs levels obtained from a medium-term model in a short-term operational model).

Investment Decision Models for Energy Storage

- Transmission losses have a relatively small effect on storage siting and investment decisions, while congestion significantly impact the model output. Congestion drives higher storage capacity siting with more of it located close to demand centers. When both forms of network constraints were introduced simultaneously, losses did not drastically change the results.
- Nevertheless, if there is no network congestion then the transmission losses play an important role to differentiate among nodes to distribute the ESS investment.
- In addition, the case studies showed that faster ramping technologies such as flywheels were favored for increased renewable penetration for their load stabilizing capabilities. However, after a certain threshold of renewable capacity was reached the investment model favored larger-scale ESS to better manage excess energy supply and avoid spillage.
- When both STESS and LTESS are considered for investment decisions, then system states and representative periods approaches are also useful to determine the optimal investment decisions. The enhanced version of representative periods (i.e., RP-TM&CI model) is the most accurate of the analyzed models. Therefore, including chronological information among representative periods may be an efficient way to include short-term variations in longer-term planning models that involve energy storage investment.
- The system states and representative periods models, as well as their enhanced versions, underestimate the ESS investment. Based on the tight

connection between VRES curtailment and storage needs, the results show that the representative periods models (RP and RP-TM&CI) have less error (both ESS investment and renewable curtailment) than the system states models (SS and SS-RFM) in the case studies that have been analyzed. We expect that these results can be replicated in more case studies, since the system states models cannot guarantee that STESS are always within bounds, leading to an overestimation of the actual energy storage capacity, and therefore, always tend to invest less in ESS than the representative periods models.

Investment Models for ESS using Power-based UC

- The case studies have shown that the proposed power-based model uses the installed investments more efficiently and more effectively, which is indicated as follows 1) it represents the reality of flexibility requirements of the power system more adequately, and 2) it adequately exploits the flexibility capabilities of the system. That is, the decisions made with the power-based model simultaneously yield lower investment costs, operating cost, CO2 emissions, and renewable curtailment with respect to the energy-based model. Although these results are obtained for a specific case study, we expect the same behavior in different case studies because the energy-based model overestimates flexibility capabilities, failing to capture the flexibility requirements such as load and VRES ramps even in a deterministic approach (i.e., without uncertainty on demand, or renewable production).
- Correctly modeling the system flexibility changes the optimal expansion capacity decisions. For instance, for the studied system, the power-based model obtains less total investment because it is more accurate in the representation of ramping characteristics for generation resources (e.g., thermal technologies and ESS), which leads to less operating cost in the real-time validation.
- The proposed semi-relaxed power-based model effectively reduces the CPU (10 to 21 times) without losing accuracy in the results compared with the non-relaxed power-based model (less than 0.2% objective function error). This is relevant for applications with large-scale long-term capacity expansion planning problems where relaxed models are more often used due to computational power limitations.
- Wrong incentives may lead to less efficient solutions. That is, forcing renewable production leads to a change in the optimal flexible re-sources the system needs, making ESS technologies with lower round-trip efficiency (such as CAES or PSH) more attractive. This general effect is model independent and applies for both energy- and power-based; how-ever, the arising optimal ESS technology mix changes drastically be-tween the energy- and power-based models. This is mainly because forcing the VRES production (through a penalization in the objective function) leads to more flexibility requirements of the thermal and ESS units, which has very different representation depending on the model. Therefore, as the flexibility requirements in the power system are higher, the differences between the energy- and power-based model in the ESS investment become larger.
- The results show an important insight for ISOs because, even without uncertainty, the current energy-based models impose more rescheduling in

the real-time operation than the power-based models. For planning authorities this is also important because decisions made with power-based models lead to a generation technology mix that is better adapted to real-time system operation.

- Although the case studies that have been analyzed in chapter 6 have not included LTESS (e.g., hydro generation), the power-based models are compatible with the enhanced version of representative periods (i.e., RP-TM&CI in chapters 3 and 5) in order to co-optimize both types of ESS including the flexibility requirements that power-based models have.

7.2 Original Contributions

Throughout this thesis, we have proposed optimization models in order to improve current operational and investment planning tools including short- and long-term operational decisions for different grid-level energy storage technologies which affects tactical and strategic planning in power systems. To summarize, the main contributions of the work presented in this thesis are the following ones, which are classified in four main areas, according to the specific objectives that have been established in this thesis, see Figure 7-1.

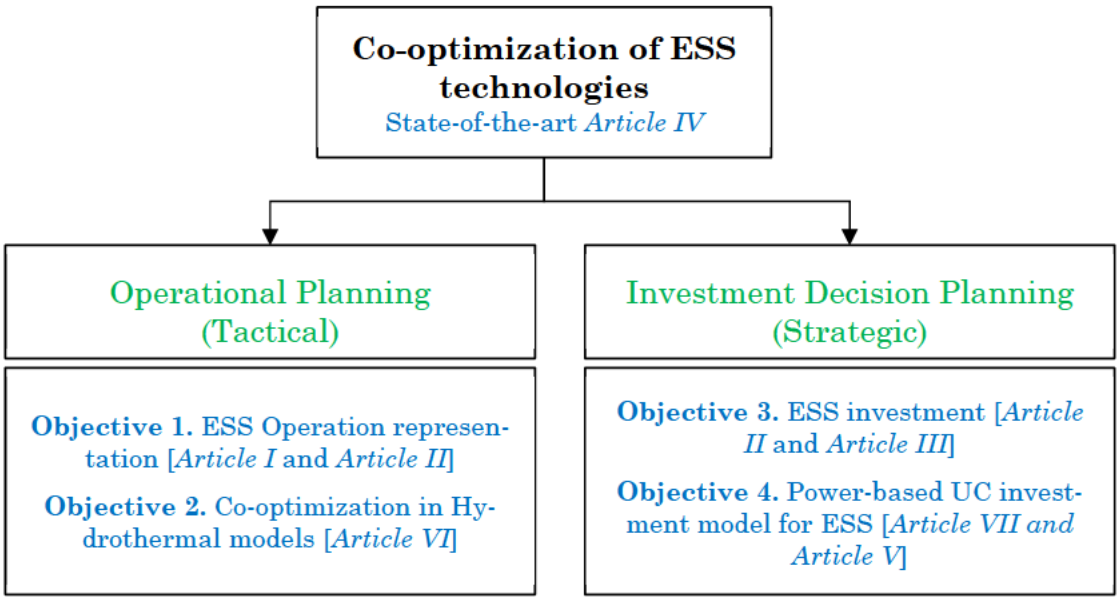


Figure 7-1. Contributions of this thesis

Journal Articles

- *Article I*. Tejada-Arango, D. A., Wogrin, S., & Centeno, E. (2018). *Representation of Storage Operations in Network-Constrained Optimization Models for Medium- and Long-Term Operation*. IEEE Transactions on Power Systems, 33(1), 386–396.
- *Article II*. Tejada-Arango, D. A., Domeshek, M., Wogrin, S., & Centeno, E. (2018). *Enhanced Representative Days and System States Modeling for Energy Storage Investment Analysis*. IEEE Transactions on Power Systems, 33(6), 6534–6544.

- *Article III*. Yacar, D., Tejada-Arango, D., & Wogrin, S. (2018). *Storage Allocation and Investment Optimization for Transmission Constrained Networks Considering Losses and High Renewable Penetration*. IET Renewable Power Generation, 12(16), 1949-1956.
- *Article IV*. Tejada-Arango, D. A., Siddiqui, A. S., Wogrin, S., & Centeno, E. (2019). *A Review of Energy Storage System Legislation in the US and the European Union*. Current Sustainable/Renewable Energy Reports, 6(1), 22-28.
- *Article V*. G. Morales-España, D.A. Tejada-Arango, *Modelling the Hidden Flexibility of Clustered Unit Commitment*. IEEE Transactions on Power Systems, pp. 1–1, 2019.

Under Review

- *Article VI*. D.A. Tejada-Arango, S. Wogrin, A. Siddiqui, E. Centeno, *Short-term storage signals in hydrothermal dispatch models using a linked representative periods approach*, October 2018. (currently under review in Energy - The International Journal).
- *Article VII*. D.A. Tejada-Arango, G. Morales-España, S. Wogrin, E. Centeno, *Power-Based Generation Expansion Planning for Flexibility Requirements*, February 2019. (currently under review in IEEE Trans-actions on Power Systems).

7.2.1 Representation of Energy Storage Operation

Article I and Article II propose improvements in current decision support models such as better dealing with short-term storage such as BESS and seasonal storage at the same time, including a network-constrained analysis. In addition, in the articles we analyze and establish the main drawbacks of the traditional modeling approaches using an hourly unit commitment model as a benchmark for the comparison of the current and proposed models. The main contributions regarding the representation of ESS operation are summarized as follows:

- The extension of the system-state optimization model for MLTOP from a single-bus system to a transmission network. This allows to determine how the congestion on the transmission network impacts the storage operation.
- The definition of an efficient clustering process for the system states considering the nodes of the network to represent storage operations. In addition, the description of the differences between the clustering process for a single node and for multiple nodes.
- The formulation of enhanced versions of system states and representative periods to preserve the chronological information of different kinds of ESS cycles (from hourly to yearly), which out-perform existing methods in terms of solution quality and CPU time and allow for the co-optimization of both short- and long-term storage operation decisions.

7.2.2 Co-optimization of ESS in hydro-thermal dispatch models

Article IV assesses the impact of short-term energy storage decisions on the opportunity cost of long-term storage through the proposal of a new optimization model for hydrothermal coordination (i.e., considering the uncertainty

in the natural hydro inflows) in which hourly opportunity costs or short-term signals are co-optimized with seasonal storage.

The main contribution is the derivation and analysis of the hourly opportunity cost of storage technologies when the proposed Linked Representative Period (LRP) model is solved. In other words, the LRP model can obtain an approximation of the ESS hourly opportunity cost within the studied time horizon without solving an hourly model. Moreover, the LRP model gets hourly opportunity cost for different types of ESS technologies which operate on different time scales (hydro vs battery). So far this was not possible because classic LDC-type models lack chronological information among the load levels.

7.2.3 Investment Decision Models for Energy Storage

Article II formulates and tests the main modeling approaches to evaluate ESS investment in power systems with high penetration of VRES by analyzing the ENTSO-E future scenarios for the Spanish case. Furthermore, Article III analyzes the influence of transmission constraints, losses, and increased VRES penetration on planning ESS allocation and investment. In this context, the main contributions are as follows:

- The formulation of enhanced versions of system states and representative periods for investment models, in order to preserve the chronological information of different kinds of ESS cycles (from hourly to yearly), which out-perform existing methods in terms of solution quality and CPU time and allow for the co-optimization investment of both short- and long-term storage.
- The comparison of system states and representative periods for ESS investment models using an hourly unit commitment model as a benchmark. This kind of comparison considering short- and long-term storage had not been developed before in the literature.
- The expansion of ESS allocation and investment model to account for both network congestion and transmission losses. A linearized approximation of ohmic losses is adapted to do this. By investigating physical constraints in a transmission network, the impact that ESS investment will have on grids operating with a high share of non-dispatchable power can be better analyzed.

7.2.4 Investment Models for ESS using Power-based UC

Article VII improves current investment models by correctly modeling power system flexibility requirements (mainly due to VRES production) that lever different ESS technologies in-vestment. Moreover, the paper compares energy-based and power-based unit commitment models and analyze the main advantages and disadvantages for the ESS investment decisions. The main contributions are summarized as follows:

- The proposal of a power-based GEP-UC model including energy storage investment that improves the classic energy-based models by representing more accurately the flexibility requirements of power systems (i.e., reserve decisions and ramping constraints). We have developed a model for ESS based on power, so it added to the power-based formulation.

- The proposal of a real-time validation stage (e.g., 5-min simulation) in order to evaluate the quality of investment and operational decisions of energy storage systems that have been obtained with the GEP-UC model.
- The proposal of a semi-relaxed version of the power-based GEP-UC model, which reduces the computational burden without losing accuracy in the results. For instance, in the proposed semi-relaxed version, the power-based GEP-UC model obtains better performance in the real-time validation stage than the traditional energy-based models, while the investment problem is solved significantly faster.

7.3 Future Research

To conclude this thesis, we summarize some interesting topics for future research which have arisen throughout this document:

Modeling improvements

- The proposed models in this thesis minimize the total operational and investment cost, which represents the point of view of a policy maker or planning authority, and it could even represent the decision-making process in a monopoly framework. However, since the liberalization wave of the electricity sector, many power systems have migrated to a competitive market framework. Moreover, operational and investment decisions of market agents depend on decisions made by other competitors in the market, as well as considering that these decisions are made in sequence (i.e., investment decisions are made first, and then operational decisions follow). Therefore, in these situations, cost minimization is not a valid assumption. From the modeling point of view, the challenge here is to adapt the proposed models in this thesis to bi-level programming models to determine investment equilibrium in competitive markets having an ESS portfolio including different kinds of storage technologies, e.g., short- and long-term energy storage systems.
- In Section 2.1, we have discussed that ESS can provide different services in power systems, such as generation, transmission, and end-user services. This thesis has focused on modeling generation services (e.g., energy arbitrage, and reserves). These types of services do not require to considering reactive power in the respective model formulations. However, some transmission services, such as voltage support, need a careful modeling of reactive power in optimization models. The main modeling challenge here is that including reactive power equations leads to a non-linear non-convex optimization problem, which is already challenging to solve without including additional ESS constraints. Therefore, two main future research lines are derived from this situation. First, AC-OPF equations can be written in three different forms: polar power-voltage (P), rectangular power-voltage (R), and rectangular current-voltage (IV). Depending on the formulation used for the AC-OPF, the non-linear terms change, e.g., P formulation has cosine and sine terms, while R and IV formulations have bi-linear terms. This is an important aspect in order to tackle the operation and investment models with AC-OPF because, depending on the AC-OPF formulation, different approaches can be applied in order to solve the problem. Thus, more research should be done in order to determine the best

formulation in the models that have been developed in this thesis. Moreover, different convexification methods for the AC-OPF formulation (e.g., Second Order Cone Programming) need to be explored in order to model reactive power constraints in medium- and long-term models. Second, reactive power characteristics of ESS depends on the type of technology. For instance, PSH has a reactive power characteristic similar to the one in a hydro generation unit, but BESS has a reactive power characteristic more similar to the one in Flexible Alternating Current Transmission System (FACTS). Thus, more research should be done in this line in order to properly model this characteristic. Considering reactive power in the proposed models in this thesis will allow to assess more accurately the ESS operation and investment in a context where they can provide multiple services.

Policy analysis

- Short-term energy storage, e.g., batteries, could provide multiple services in power systems such as: energy arbitrage, renewable support, reserve markets, security of supply, etc.; and consequently, new ways to consider these multiple services in the planning tools or decision support models are needed. Nevertheless, proposed models in this thesis may be an option to carry out exhaustive analysis focusing on the added value for society of storing electricity by simultaneously providing multiple services.
- In addition, the proposed models in this thesis may also be useful to explore the economic viability of the maximum social benefit solutions for private investors in ESS. They may also help to determine the gap between the solutions of maximum social benefit (what we would like to happen) and maximum benefit for private investors (what will happen).

A OPERATION AND INVESTMENT FORMULATIONS

A.1 Notation

In the following formulation "p/s" refer to the parameters used to identify time divisions: periods (e.g., 1 h) in the detailed model and states in the system states model respectively.

Indices and Sets

$p \in \mathrm{P}$	Periods (hours)
$p_l(p)$	Subset with the last period of the time horizon
$s, s' \in S$	System states
$k \in \mathrm{K}$	Periods in which storage limit are imposed in system states
$g \in G$	Generation units (thermal or storage)
$t(g)$	Subset of thermal generation units
$h(g)$	Subset of storage units
$h_l(g)$	Subset of long-term storage (e.g., hydro) units
$h_s(g)$	Subset of short-term storage (e.g., batteries) units
$n, n' \in \mathrm{N}$	Electrical nodes o buses
$n_s(n)$	Subset of electrical nodes or buses without slack bus
c	Circuits
$\mathcal{G}_{gn}$	Generators g connected to bus n
$\Theta_{nn'c}$	Circuits c connected between bus n' and n
$rp \in RP$	Set of representative periods (e.g., days, weeks)
Γ_{rpp}	Injective map of each period p to a representative period rp
$H_{pp'}$	Injective map of each period p to a period $p' \in \Gamma_{rpp}$
$p_f(p, rp)$	Subset with the first period p of the representative period rp

Parameters

C_g^{fuel}	Cost of consumed fuel [k€/MJ]
α_g	Variable term of fuel consumption [MJ/MWh]
β_g	Fixed term of fuel consumption [MJ]
γ_g	Fuel consumption during the startup [MJ]
C_g^{om}	Cost of operation and maintenance [k€/MWh]
$D_{p/s\,n}$	Electricity demand per node [MW]
$V_{p/s\,n}^{max}$	Renewable production per node (e.g., wind or solar) [MW]
Q_g^{max},Q_g^{min}	Upper and lower bound on production [MW]
SRR_g	Maximum 10-minute ramp [MW]
X^{res}	Operating reserve [p.u.]
$W0_h$	Initial storage level [MWh]
W_h^{max},W_h^{min}	Upper and lower bound on energy storage [MWh]
W_h^{fin}	Minimum final storage level [MWh]
$I_{p/s\,h}$	Hourly energy inflows [MWh]
η_h	Efficiency of storage unit [p.u.]
B_h^{max}	Upper bound on charging/pumping [MW]
T_s	Duration of state [h]
$TC_{n\,n'c}^{max}$	Transmission capacity of circuit c [MW]

$ISF_{n\,n'c\,n_s}$	Injection Shift Factors [p.u.]
$N_{s\,s'}$	Transition matrix between states
$F_{s\,s'k}$	Frequency matrix between states and changes
$RFM_{s\,s'k}$	Reduced Frequency Matrix between states and changes
WG_{rp}	Weight of representative periods [h]
$NRP_{rp\,rp'}$	Transition matrix between representative periods
NP_{rp}	Number of periods at each representative period [h]
M	Moving window for storage level [h]
C_h^{inv}	Investment cost for storage units [€/MW]
EPR_h^{max} EPR_h^{min}	Maximum and minimum energy to power ratio [h]

Variables

$q_{p/s\,g}$	Power production [MW]
$\hat{q}_{p/s\,g}$	Power production above Q_g^{min} [MW]
$v_{p/s\,n}$	Renewable production [MW]
$r_{p/s\,g}$	Spinning reserve [MW]
$w_{p/s\,h}$	Storage level [MWh]
$\Delta w_{s\,s'\,h}$	Difference in storage [MWh]
$b_{p/s\,h}$	Hourly charged/pumped power [MW]
$sp_{p/s\,h}$	Hourly energy spillage [MWh]
$pf_{p/s\,nn'c}$	Power flow per circuit [MW]
$pns_{p/s\,n}$	Power not supply per node [MW]
$u_{p/s\,g}$	Binary dispatch decision [0-1]
$y_{p/s\,g}$	Binary startup decision [0-1]
$y_{s\,s'\,g}$	Binary startup decision for state model [0-1]
x_h	Storage investment [MW]

A.2 Hourly Model

The following equations describe the hourly unit commitment model used as the benchmark to test the proposed models in this thesis.

$$min_{\Omega} \sum_{p,t}\left\{C_t^{fuel}\cdot\left[\beta_t u_{pt}+\gamma_t y_{pt}+\alpha_t q_{pt}\right]+C_t^{om}q_{pt}\right\}+\sum_h C_h^{inv}x_h \qquad \text{(A.2-1)}$$

Subject to:

$$\sum_{t\in\mathcal{G}} q_{pt}+\sum_{h\in\mathcal{G}}\left(q_{ph}-b_{ph}\right)+v_{pn}+\sum_{n'c\in\Theta}\left(pf_{pn'nc}-pf_{pnn'c}\right)+pns_{pn} = D_{pn} \quad \forall\, p,n \qquad \text{(A.2-2)}$$

$$pf_{pnn'c} = \sum_{n_s} ISF_{nn'cn_s} \cdot \left[\sum_{t \in \mathcal{G}_{tn_s}} q_{pt} + \sum_{h \in \mathcal{G}_{hn_s}} \left(q_{ph} - b_{ph} \right) + v_{pn_s} + pns_{pn_s} - D_{pn_s} \right] \quad \forall\, nn'c \in \Theta\,, p \tag{A.2-3}$$

$$q_{pt} = Q_t^{min} u_{pt} + \hat{q}_{pt} \qquad \forall\, p,t \qquad \text{(A.2-4)}$$

$$0 \le \hat{q}_{pt} \le \left(Q_t^{max} - Q_t^{min} \right) u_{pt} \qquad \forall\, p,t \qquad \text{(A.2-5)}$$

$$u_{pt} - u_{p-1,t} \le y_{pt} \qquad \forall\, p,t \qquad \text{(A.2-6)}$$

$$r_{pt} + q_{pt} \le u_{pt} Q_t^{max} \qquad \forall\, p,t \qquad \text{(A.2-7)}$$

$$0 \le r_{pt} \le SRR_t \qquad \forall\, p,t \qquad \text{(A.2-8)}$$

$$\sum_t r_{pt} \ge X^{res} \cdot \sum_n D_{pn} \qquad \forall\, p \qquad \text{(A.2-9)}$$

$$u_{pt}, y_{pt} \in \{0,1\} \qquad \forall\, p,t \qquad \text{(A.2-10)}$$

$$w_{ph} = w_{p-1,h} + W0_{p=1,h} + I_{ph} - q_{ph} - sp_{ph} + \eta_h b_{ph} \qquad \forall\, p,h \qquad \text{(A.2-11)}$$

$$0 \le v_{pn} \le V_{pn}^{max} \qquad \forall\, p,n \qquad \text{(A.2-12)}$$

$$0 \le q_{ph} \le Q_h^{max} + x_h \qquad \forall\, p,h \qquad \text{(A.2-13)}$$

$$0 \le b_{ph} \le B_h^{max} + \eta_h x_h \qquad \forall\, p,h \qquad \text{(A.2-14)}$$

$$0 \le sp_{ph} \qquad \forall\, p,h \qquad \text{(A.2-15)}$$

$$\left| pf_{pnn'c} \right| \le TC_{nn'c}^{max} \qquad \forall\, \Theta_{nn'c}, p \qquad \text{(A.2-16)}$$

$$W_h^{min} + EPR_h^{min} x_h \le w_{ph} \le W_h^{max} + EPR_h^{max} x_h \qquad \forall\, p,h \qquad \text{(A.2-17)}$$

$$w_{p_l,h} \ge W_h^{fin} \qquad \forall\, h \qquad \text{(A.2-18)}$$

The objective function (A. 2-1) minimizes storage investment costs and the total operating cost of the system (e.g. startup costs, fixed costs, variable costs, operations and maintenance costs, and penalties for spillage and energy not supplied). Constraint (A.2-2) is the demand balance equation. Constraint (A.2-3) represents the power flow equation using Injection Shift Factors (ISF). Constraints (A.2-4) and (A.2-5) ensure thermal unit production is within minimum and maximum capacity. Constraint (A.2-6) is the startup constraint of the unit-commitment. Constraints (A.2-7) to (A.2-9) are reserve constraints. Constraint (A.2-10) states that the commitment and connection variables are binary. Constraint (A.2-11) is the storage constraint which states that the storage in any hour is the storage in the previous hour plus the net charging and discharging in the current hour. Constraints (A.2-12) to (A.2-17) keep within bounds the renewable production per node, the power output per storage unit, the pumped power per storage unit, the energy spillage, the power flow through a line, and the amount of energy stored in each storage unit. Constraints (A.2-13) and (A.2-14) include the power capacity increase due to the storage investment variable. In models to evaluate only ESS operation,

the investment variable x_h is set to zero. Constraint (A.2-17) includes the energy capacity increase considering parameters EPR_h^{max} and EPR_h^{min}. These parameters describe the relationship between the energy that can be stored (i.e., maximum and minimum respectively) and the nominal power of the equipment. Finally, constraint (A.2-18) establishes the minimum storage level at the last period of the time horizon.

A.3 System States Model

This section presents the formulation of the system states model as conceived in [94].

$$min_{\Omega} \sum_{s,t}\left\{C_t^{fuel}\cdot\left[T_s\beta_t u_{st} + \sum_{s'\neq s} N_{s's}\gamma_t y_{s'st} + T_s\alpha_t q_{st}\right] + C_t^{om}T_s q_{st}\right\} + \sum\nolimits_h C_h^{inv}x_h \quad \text{(A.3-1)}$$

Subject to:

$$\sum_{t\in\mathcal{G}} q_{st} + \sum_{h\in\mathcal{G}}(q_{sh}-b_{sh}) + v_{sn} + \sum_{n'c\in\Theta}(pf_{sn'nc} - pf_{snn'c}) + pns_{sn} = D_{sn} \quad \forall\, s,n \quad \text{(A.3-2)}$$

$$pf_{snn'c} = \sum_{n_s} ISF_{nn'cn_s}\cdot\left[\sum_{t\in\mathcal{G}_{tn_s}} q_{st} + \sum_{h\in\mathcal{G}_{hn_s}}(q_{sh}-b_{sh}) + v_{sn_s} + pns_{sn_s} - D_{sn_s}\right] \quad \forall\, nn'c\in\Theta, s \quad \text{(A.3-3)}$$

$$q_{st} = Q_t^{min}u_{st} + \hat{q}_{st} \qquad \forall\, s,t \qquad \text{(A.3-4)}$$

$$0 \le \hat{q}_{st} \le \left(Q_t^{max} - Q_t^{min}\right)u_{st} \qquad \forall\, s,t \qquad \text{(A.3-5)}$$

$$u_{st} - u_{s',t} \le y_{s'st} \qquad \forall\, s,t \qquad \text{(A.3-6)}$$

$$r_{st} + q_{st} \le u_{st}Q_t^{max} \qquad \forall\, s,t \qquad \text{(A.3-7)}$$

$$0 \le r_{st} \le SRR_t \qquad \forall\, s,t \qquad \text{(A.3-8)}$$

$$\sum_t r_{st} \ge X^{res}\cdot\sum_n D_{sn} \qquad \forall\, s \qquad \text{(A.3-9)}$$

$$u_{st}, y_{s'st} \in \{0,1\} \qquad \forall\, s,t \qquad \text{(A.3-10)}$$

$$0 \le v_{sn} \le V_{sn}^{max} \qquad \forall\, s,n \qquad \text{(A.3-11)}$$

$$0 \le q_{sh} \le Q_h^{max} + x_h \qquad \forall\, s,h \qquad \text{(A.3-12)}$$

$$0 \le b_{sh} \le B_h^{max} + \eta_h x_h \qquad \forall\, s,h \qquad \text{(A.3-13)}$$

$$0 \le sp_{sh} \qquad \forall\, s,h \qquad \text{(A.3-14)}$$

$$|pf_{snn'c}| \le TC_{nn'c}^{max} \qquad \forall\, \Theta_{nn'c}, s \qquad \text{(A.3-15)}$$

$$\Delta w_{ss'h} = 0.5\cdot(I_{sh} + I_{s'h} + \eta_h b_{sh} + \eta_h b_{s'h} - q_{sh} - q_{s'h} - sp_{sh} - sp_{s'h}) \qquad \forall\, s,s',h \qquad \text{(A.3-16)}$$

$$\sum_{\substack{s,s' \text{ s.t.} \\ N_{ss'}>0}} N_{ss'} \cdot \Delta w_{ss'h} \geq W_h^{fin} - W0_h + EPR_h^{min} x_h \qquad \forall\, h \qquad \text{(A.3-17)}$$

$$\sum_{\substack{s,s' \text{ s.t.} \\ N_{ss'}>0}} N_{ss'} \cdot \Delta w_{ss'h} \leq W_h^{max} - W0_h + EPR_h^{max} x_h \qquad \forall\, h \qquad \text{(A.3-18)}$$

$$\sum_{\substack{s,s' \text{ s.t.} \\ F_{ss'k}>0}} F_{ss'k} \cdot \Delta w_{ss'h} \geq W_h^{min} - W0_h + EPR_h^{min} x_h \qquad \forall\, h,k \qquad \text{(A.3-19)}$$

$$\sum_{\substack{s,s' \text{ s.t.} \\ F_{ss'k}>0}} F_{ss'k} \cdot \Delta w_{ss'h} \leq W_h^{max} - W0_h + EPR_h^{max} x_h \qquad \forall\, h,k \qquad \text{(A.3-20)}$$

The objective function (A.3-1) incorporates storage investment and operational costs just as in the hourly model. The costs of each state are weighted by the number of hours in the time horizon that belong to that state, and the startup costs are multiplied by the transition matrix which gives the number of transitions between each set of states. Constraints (A.3-2) to (A.3-15) are formulated exactly as in the hourly model in previous section, except that they are defined for each system state 's' rather than each hour 'p'. Constraints (A.3-16) to (A.3-20) are the system states formulation of the storage constraints. These constraints are explained in more detailed due to their relevance in this model.

Storage Representation in the System States Model

In order to properly represent storage units in the system state approach, we must first define a new variable $\Delta w_{ss'h}$. This variable represents the difference in storage level for a specific storage unit h between two different states. Solving the model with this variable can allow us to analyze storage level differences as well as charge, discharge and spillage information. This information can then be backtracked to obtain the corresponding hourly storage levels on storage units. Mathematically, this variable is defined as the central finite difference between two states s and s' in constraint (A.3-16). Therefore, the storage level difference variable is calculated by taking the average of the sum of the difference in storage level for one of the hour blocks in one state s (charge minus discharge minus spillage plus inflows) to the storage level difference in another hour block in a different state s'. The sum of the series is $\Delta w_{ss'h}$ multiplied by the time that transition occurs, value included in the transition matrix N'_{ss}, results in the evolution of the storage level from beginning to end of time horizon. Therefore, constraint (A.3-17) defines the lower boundary of the overall storage level difference from the initial and final levels, and constraint (A.3-18) is the upper bound of the overall storage level difference between the maximum and initial storage levels for a storage unit.

Until now, the storage level along the time horizon is not guaranteed to be within bounds. More constraints need to be included to make sure that the storage level is limited to intermediate (between the first and the last hour of the time horizon) hours. To do this index k is used. It denotes a set of hours at which the imposed storage level limitations still hold. Another parameter,

$F_{ss'k}$ is also introduced into the model. This parameter represents the frequency with which a storage unit moves between hours of state s to hours of state s' before hour k. The lower and upper storage boundaries for all k are imposed in constraints (A.3-19) and (A.3-20) respectively. In order to keep the system states approximation computationally tractable and efficient, it is crucial that the value of k stays relatively small in comparison to the number of hours that will be reduced. A naïve approach is for example selecting 10% of the total hours, however, the Iterative Approach, which is explained latter in this section, goes into further detail on the process needed to obtain a small value of k and the corresponding frequency matrix $F_{ss'k}$. It is important to note that once the system state model is solved, the results can be used to calculate an equivalent hourly solution. This ex post facto solution is made possible by the fact that the clustering process uniquely assigns each hour to one state.

It is important to point out that without the system state formulation presented in this section, the storage level cannot be properly modeled in an LDC approach with a network. In fact, the LDC approach traditionally only considers the storage balance constraint in aggregated periods (e.g., weeks, or months). Therefore, it is not possible to obtain the storage level at block (and subsequently hourly) level. For hydro reservoirs with storage capacity higher than a week or month this approach is enough. However, storage units with lower storage capacity (e.g., batteries or daily reservoirs) cannot be properly modeled in the LDC approach. The modeling of these types of storage facilities is becoming more relevant due to the increasing penetration of renewable energy in power systems. Then, the system state framework gives us a way to approximate the future operation of storage units for MLTOP no matter the storage capacity of the facility.

Iterative Approach for the Formulation of Lower and Upper Storage Level Bounds

Much like the ability to calculate hourly solutions, obtaining the set of hours K is also based on the fact that each hour is assigned to only one system state. Therefore, the set K represents the set of hours at which the upper and lower bound constraints (A.3-19) and (A.3-20) for the storage unit are imposed. A naïve definition of set K is the hours where there is a change from one state to another in the hourly list of the state assigned to each hour. $F_{ss'k}$ represents the number of times we have changed from one state to another until time step k. For instance, the transmission matrix $N_{ss'}$ represents all transitions between states s and state s' over all the time horizon, while the frequency matrix $F_{ss'k}$ represents the amount of transitions between states s and state s' until hour (time step) k. In fact, if k is equal to the last hour then both matrices have the same values at each element. Therefore, every element of $F_{ss'k}$ must be lower or equal to the same element at $N_{ss'}$. In other words, it can be said that $F_{ss'k}$ is defined similarly to the transmission matrix $N_{ss'}$ at each time step k.

As mentioned previously the computational efficiency of the storage approximation in the system state model depends heavily on the cardinality of set K. Therefore, reducing the number of hours for which storage level bounds are explicitly defined is favorable. In [51] a heuristic method is used in order to

reduce the size of set K and the corresponding frequency matrix $F_{ss'k}$ in a single-node case study. The heuristic approach is based on the idea of only enforcing upper and lower bounds on the storage level when there is a change in operations from charging to discharging. The drawback of this strategy is that it becomes necessary first to identify the charging and discharging times from the system's historical data. This makes the limiting assumption that past charge/discharge behavior will be representative of future actions, failing to take into consideration changes in system components (e.g. new lines or generators).

An alternative iterative approach, that is described in [151] can also be used to define the hours associated with set K. In this thesis, the iterative approach has been applied to a transmission network problem, and it is shown that it is still valid for multiple nodes. This iterative approach is called Incremental Bounding Algorithm (IBA). The first step required for solving via this method is to solve the system states model without constraints (A.3-19) and (A.3-20), which means without limits on the storage level. Therefore, the post factum hourly storage level may fall outside the bounds at specific hours. The hours for which the storage level is out of bounds are added to set K and the corresponding frequency matrix $F_{ss'k}$ is calculated. The system state model is then solved once again but this time including constraints (A.3-19) and (A.3-20). Once again, the hourly storage levels must be calculated ex post in order to verify that they all fall within the boundaries. If this condition is not satisfied, the process is repeated until there are no more hours that must be added to set K in order to properly maintain boundaries. This iterative process helps ensure the smallest k value possible, reducing the computational burden of the system state model. Therefore, this iterative approach is applied in this paper in order to obtain set K and frequency matrix $F_{ss'k}$. This iterative process is explained through the following pseudocode, where ES(p) stands for energy storage at a given hour p, and UB and LB stands for upper and lower bound respectively.

```
{k} = ∅              // K set starts empty
SOLVE system states model
p = 1                // Set start to first hour
While (ES(p to end) > UB or ES(p to end) < LB)
  {k} = {k} U {p}  // Add c hour when exceeds bound
  SOLVE system states model    // Solve model again
```

If a value of energy storage is higher than the upper bound the while loop is entered. If the energy stored of an hour is below the feasible minimum, lower bound, the while loop is entered [151].

A.4 Enhanced System States Model

The objective function in the enhanced system states model is the same as in the former system states model (A.3-1). In addition, constraints (A.3-2) to (A.3-18) remain also without changes.

The difference between the two models lies in the handling of storage which has been separated into long- and short-term storage, each with its own set

of constraints. (A.4-1) and (A.4-2) take the same form as (A.3-19) and (A.3-20), but are only applied to long-term storage, which is likely to go through only one or two cycles per year. Set k is a subset of hours in the time horizon in which the upper and lower bound are checked. At each hour k, (A.3-19) and (A.3-20) use the frequency matrices to add up all changes in storage from the beginning of the time horizon to hour k and check that the total is within bounds. (A.4-3) and (A.4-4), represent the storage constraints for short-term storage. At each hour k, they add up all the net changes in storage since the last hour k and constrain that sum to be within bounds. This is done with the aid of the Reduced Frequency Matrix (RFM), an innovation of this model which is just the difference between the frequency matrix ($F_{ss'k}$) corresponding to the current hour k and that corresponding to the previous element in set k, that is, $k-1$. In other words, the difference between these two elements or hours in the set k could be understood as a moving window. It is important to mention that despite the use of the RFM, the storage level could be out of bounds because the hours in set k are predefined in the model and we do not know in advance the storage level value at each hour in set k. The best practice for reducing the number of hours in which the storage levels can be out of bounds is to predefine the moving window considering the smallest storage cycle in the power system.

$$\sum_{\substack{s,s' \ s.t. \\ F_{ss'k}>0}} F_{ss'k} \cdot \Delta w_{ss'h_l} \geq W_{h_l}^{min} - W0_{h_l} + EPR_h^{min} x_h \qquad \forall\, h_l, k \quad \text{(A.4-1)}$$

$$\sum_{\substack{s,s' \ s.t. \\ F_{ss'k}>0}} F_{ss'k} \cdot \Delta w_{ss'h_l} \leq W_{h_l}^{max} - W0_{h_l} + EPR_h^{max} x_h \qquad \forall\, h_l, k \quad \text{(A.4-2)}$$

$$\sum_{\substack{s,s' \ s.t. \\ RFM_{ss'k}>0}} RFM_{ss'k} \cdot \Delta w_{ss'h_s} \geq W_{h_s}^{min} - W0_{h_s} + EPR_h^{min} x_h \qquad \forall\, h_s, k \quad \text{(A.4-3)}$$

$$\sum_{\substack{s,s' \ s.t. \\ RFM_{ss'k}>0}} RFM_{ss'k} \cdot \Delta w_{ss'h_s} \leq W_{h_s}^{max} - W0_{h_s} + EPR_h^{max} x_h \qquad \forall\, h_s, k \quad \text{(A.4-4)}$$

A.5 Representative Periods Model

This section describes the RP model which is a commonly used method of reducing temporal information. Although the model is general enough to work with RPs of any length, we will speak of representative days for the sake of simplicity. The formulation is roughly the same as that of the hourly model, except the constraints only apply to the hours within the representative days.

$$min_{\Omega} \sum_{p,rp \epsilon \Gamma_{rpp}} \left\{ WG_{rp} \cdot \sum_t \left\{ C_t^{fuel} \cdot \left[\beta_t u_{pt} + \gamma_t y_{pt} + \alpha_t q_{pt} \right] + C_t^{om} q_{pt} \right\} \right\} + \sum_h C_h^{inv} x_h \qquad \text{(A.5-1)}$$

Subject to:

Equations (A.2-2) to (A.2-18) $\forall\, p \epsilon \Gamma_{rpp}$

$$w_{p=p_f(p,rp)+NP_{rp}-1,h} \geq w_{p=p_f(p,rp),h} \qquad \forall\,(p,rp)\epsilon\Gamma_{rpp},h \qquad \text{(A.5-2)}$$

The objective function (A.5-1) minimizes the storage investment cost and operational cost just as in the hourly model, except that the operational costs associated with each day are multiplied by the number of days in the time horizon that are represented by it to yield the cost for the entire time horizon. The RP model is constrained to equations (A.2-2) to (A.2-18) from the HM benchmark model. Nevertheless, in the RP model, equations (A.2-2) to (A.2-18) only apply to hours belonging to the selected representative days.

Equation (A.5-2) is a special constraint introduced into the RP model that guarantees that the amount of energy stored in each unit at the end of each representative day is greater than or equal to the amount of energy in storage at the beginning of the day. Since each day is calculated separately, this prevents a unit from finishing a day with less energy than the starting level of the next day, and thus creating energy from nothing. This is a very simple way to deal with the maximum energy storage per year. Other approaches ensure that the change accumulated over each representative period does not exceed the storage limits and ensure balance over the whole year. However, for the sake of simplicity, these types of approaches are not analyzed in this paper.

Despite the incorporation of (A.5-2), each representative day is independent of the others and the RP model does not guarantee chronological continuity among the representative days for the ESS.

A.6 Enhanced Representative Periods Model

This section shows the Representative Period with Transition Matrix and Cluster Indices (RP-TM&CI) model which is the second original contribution of this paper. Although the model is sufficiently general to be able to work with representative periods of any length, we will once again speak of representative days for the sake of simplicity.

The objective function has the same formulation as the regular representative day model, i.e., (A.5-1). The RP-TM&CI model is constrained with equations (A.2-2) to (A.2-18) for all the hours belonging to the selected representative days. Constraint (A.6-1) is an innovation of this model. It creates continuity between the representative days and prevents unnecessary startups by using a transition matrix to require that for any pair of representative days that transition from one to the other, the thermal units that are on in the last hour of the first are also on in the first hour of the second. As written here, if there is even one transition between the two days, this constraint is applied. However, the constraint could be set to take effect only if there is a considerable number of transitions between the two days, 5 or 10% of the transitions in the time horizon, for example. Constraint (A.6-2) is the second innovation of this model; it creates the continuity in storage across the entire time horizon that allows for the modeling of long-term storage. It does this by checking at regular intervals (1 week) that all the energy charged and discharged since the previous week plus the total energy at the last check point are within bounds. This is possible because, as a result of the clustering procedure to determine the representative days, we know the Cluster Indices (CI), which

is a numeric column vector where each row indicates the cluster assignment (i.e. representative day) of the corresponding day of the year. This information is included in the model using the subset $H_{pp'}$.

$$u_{p'=p_f(p',rp')+NP_{rp}-1,t} = u_{p=p_f(p,rp),t} \quad \forall\, t,(p,rp)\epsilon\Gamma_{rpp},rp' \,/\, NRP_{rp\,rp'} > 0 \tag{A.6-1}$$

$$w_{ph} = w_{p-M,h} + W0_{p=1,h} + \sum_{p'=p-M+1}^{p} \sum_{p''\in H_{p'p''}} \left(I_{p''h} - q_{p''h} - sp_{p''h} + \eta_h b_{p''h}\right) \quad \forall\, p,h \tag{A.6-2}$$

B HYDROTHERMAL DISPATCH FORMULATIONS

B.1 Notation

In the following formulation, "$*$" refers to the parameters or variables used to identify time divisions: p for hours in the detailed hourly model, (m, w, l) in the load-levels model, and (rp, k) in the linked representative periods model respectively.

Indices and Sets

$p \in \mathcal{P}$	Periods (e.g., hours)
$m \in \mathcal{M}$	Aggregation of hours (e.g., months)
$MP_{m,p}$	Relation among hours and months
$w \in \mathcal{W}$	Aggregation of load levels (e.g., weekdays)
$l \in \mathcal{L}$	Load levels or load blocks
$rp \in \mathcal{RP}$	Representative periods (e.g., days)
$TM_{rp',rp}$	Transitions among rp
$k \in \mathcal{K}$	Hours inside a representative period
$CI_{p,rp,k}$	Cluster index
$r \in \mathcal{R}$	Reservoirs
$g \in \mathcal{G}$	Generators
$t \subset g$	Thermal units
$s \subset g$	Storage units
$b \subset s$	Short-term storage units (e.g., batteries)
$h \subset s$	Hydro units
$HUR_{h,r}$	Hydro plant upstream of reservoir
$HPR_{h,r}$	Pumped hydro plant upstream of reservoir
$RUH_{r,h}$	Reservoir upstream of hydro plant
$RPH_{r,h}$	Reservoir upstream of pumped hydro plant
$RUR_{r,r}$	Reservoir upstream of reservoir
$\omega \in \Omega$	Scenarios
$a(\omega)$	Scenario tree relations

Parameters

d_*, o_*	Demand, operating reserve [MW]
wg_*	Load level duration or rp weight [h]
$\bar{p}_g, \underline{p}_g$	Maximum, minimum output [MW]
f_t, v_t	No load cost [\$/h], variable cost [\$/MWh]
su_t, sd_t	Startup, shutdown cost [\$]
c_h, η_s	Production function and efficiency [p.u.]
$\bar{r}_r, \underline{r}_r$	Maximum and minimum reserve [hm^3]
r'_r	Initial and final reserve [hm^3]
$\overline{soc}_b, \underline{soc}_b$	Maximum, minimum state of charge [p.u.]
$i^{\omega}_{*,r}$	Stochastic hydro inflows [m^3/s]
p^{ω}_*	Scenario probability [p.u.]
v'	Energy not served cost [\$/MWh]
v''	Operating reserve not served cost [\$/MWh]

Variables

$UC^{\omega}_{*,t}, SU^{\omega}_{*,t}, SD^{\omega}_{*,t}$	Commitment, startup, and shutdown {0,1}
$P^{\omega}_{*,g}$	Production of generation units [MW]
$P'^{\omega}_{*,t}$	Production above minimum output [MW]
$C^{\omega}_{*,h}, C^{\omega}_{*,s}$	Consumption of a hydro/storage unit [MW]
$O^{\omega}_{*,g}$	Operating reserve of generation unit
$R^{intra,\omega}_{*,r}, R^{inter,\omega}_{*,r}$	Intra and inter reservoir level [hm^3]
$S^{\omega}_{*,r}$	Reservoir spillage [hm^3]
$ENS^{\omega}_{*}, RNS^{\omega}_{*}$	Energy and reserve not served [MW]
$SoC^{intra,\omega}_{*,b}, SoC^{inter,\omega}_{*,b}$	State-of-charge of a BESS [p.u.]

B.2 Hourly Stochastic Hydrothermal Dispatch

The following equations describe the hourly UC model used as the benchmark to test the proposed models.

$$\min \sum_{\omega pt} p^{\omega}_{p} \cdot su_t \cdot SU^{\omega}_{pt} + \sum_{\omega pt} p^{\omega}_{p} \cdot sd_t \cdot SD^{\omega}_{pt} + \sum_{\omega pt} p^{\omega}_{p} \cdot f_t \cdot UC^{\omega}_{pt} + \sum_{\omega pt} p^{\omega}_{p} \cdot v_t \cdot P^{\omega}_{pt} + \sum_{\omega p} p^{\omega}_{p} \cdot v' \cdot ENS^{\omega}_{p} + \sum_{\omega p} p^{\omega}_{p} \cdot v'' \cdot RNS^{\omega}_{p} \quad \text{(B.2-1)}$$

Subject to:

$$\sum_{g} O^{\omega}_{pg} + RNS^{\omega}_{p} \geq o_p \ \forall \omega p \quad \text{(B.2-2)}$$

$$\sum_{g} P^{\omega}_{pg} - \sum_{s} \frac{C^{\omega}_{ps}}{\eta_s} + ENS^{\omega}_{p} = d_p \ \forall \omega p \quad \text{(B.2-3)}$$

$$UC^{\omega}_{pt} - UC^{\omega'}_{p-1t} = SU^{\omega}_{pt} - SD^{\omega}_{pt} \ \forall \omega pt \ \omega' \epsilon \ a(\omega) \quad \text{(B.2-4)}$$

$$P'^{\omega}_{pt} + O^{\omega}_{pt} \leq \left(\bar{p}_t - \underline{p}_t\right) \cdot \left(UC^{\omega}_{pt} - SU^{\omega}_{pt}\right) \forall \omega pt \quad \text{(B.2-5)}$$

$$P'^{\omega}_{pt} + O^{\omega}_{pt} \leq \left(\bar{p}_t - \underline{p}_t\right) \cdot \left(UC^{\omega}_{pt} - SD^{\omega}_{p+1t}\right) \forall \omega pt \quad \text{(B.2-6)}$$

$$P^{\omega}_{pt} = \underline{p}_t \cdot UC^{\omega}_{pt} + P'^{\omega}_{pt} \quad \forall \omega pt \quad \text{(B.2-7)}$$

$$P^{\omega}_{ph} \leq \bar{p}_h \quad \forall \omega ph \quad \text{(B.2-8)}$$

$$R_{p-1r}^{intra,\omega'} - R_{pr}^{intra,\omega} + i_{pr}^{\omega} - S_{pr}^{\omega} + \underbrace{\sum_{r' \in RUR_{r',r}} S_{pr'}^{\omega}}_{Water\ spillage\ from\ upstream\ reservoirs} + \underbrace{\sum_{h \in HUR_{h,r}} {}^{P_{ph}^{\omega}}/_{c_h}}_{Turbined\ water\ from\ upstream\ hydro\ plants} - \underbrace{\sum_{h \in RUH_{r,h}} {}^{P_{ph}^{\omega}}/_{c_h}}_{Turbined\ \ water\ from\ hydro\ plants\ in\ reservoir\ r} + \underbrace{\sum_{h \in HPR_{h,r}} {}^{C_{ph}^{\omega}}/_{c_h}}_{Pumped\ water\ from\ hydro\ plants\ to\ reservoir\ r} - \underbrace{\sum_{h \in RPH_{r,h}} {}^{C_{ph}^{\omega}}/_{c_h}}_{Pumped\ water\ to\ other\ reservoirs} = 0 \quad \forall \omega pr \ \ \omega' \epsilon\, a(\omega) \tag{B.2-9}$$

$$SoC_{p-1b}^{intra,\omega'} - SoC_{pb}^{intra,\omega} - P_{pb}^{\omega} + C_{pb}^{\omega} = 0 \quad \forall \omega pb \quad \omega' \epsilon\, a(\omega) \tag{B.2-10}$$

$$R_{|P|r}^{intra,\omega} = r_r' \quad \forall \omega r \tag{B.2-11}$$

$$\underline{r}_r \leq R_{pr}^{intra,\omega} \leq \overline{r}_r \quad \forall \omega pr \tag{B.2-12}$$

$$\underline{soc}_b \leq SoC_{pb}^{intra,\omega} \leq \overline{soc}_b \quad \forall \omega pb \tag{B.2-13}$$

$$0 \leq P_{pg}^{\omega} \leq \overline{p}_g \quad \forall \omega pg \tag{B.2-14}$$

$$0 \leq C_{ps}^{\omega} \leq \overline{p}_s \quad \forall \omega ps \tag{B.2-15}$$

$$UC_{pt}^{\omega}, SU_{pt}^{\omega}, SU_{pt}^{\omega} \in \{0,1\} \quad \forall \omega pt \tag{B.2-16}$$

The objective function in (B.2-1) minimizes the expected value of operational cost. Constraints (B.2-2) and (B.2-3) represent the operating reserve and the demand-balance equations respectively. Equation (B.2-4) is the logical relationship among the binary variables for UC. Constraints (B.2-5) to (B.2-7) ensure thermal unit production is within minimum and maximum capacity, while (B.2-8) ensures it for hydro units. Also, (B.2-9) defines the water balance for each reservoir considering the hydro topology. Constraint (B.2-10) defines the state of charge for each short-term storage (e.g., batteries). Notice that constraints (B.2-9) and (B.2-10) are equivalent if, for example, a hydro reservoir has a pump unit which is not in a hydro basin and it has no hydro inflows. However, we keep both constraints in order to facilitate the distinction between both types of storage technologies. Constraints (B.2-9) and (B.2-10) also are stated for $\omega' \epsilon\, a(\omega)$, which allows to relate the different scenarios through the scenario tree. Constraint (B.2-11) establishes the final reservoir level at the last period of the time horizon. Equations (B.2-12) to (B.2-15) maintain bounds for the reservoir level, the state of charge, the power output, and the charged power per storage unit. Finally, (B.2-16) states that the commitment and connection variables are binary.

B.3 LDC Stochastic Hydrothermal Dispatch

This section shows the load levels formulation based on the demo version of StarGen Lite model (Medium Term Stochastic Hydrothermal Coordination Model) [20], which is used for modeling hydrothermal dispatch models in [16].

$$\begin{aligned}\min \sum_{\omega mwt} p_m^{\omega} \cdot su_t \cdot SU_{mwt}^{\omega} + \sum_{\omega mwt} p_m^{\omega} \cdot sd_t \cdot SD_{mwt}^{\omega} \\ + \sum_{\omega mwlt} p_m^{\omega} \cdot wg_{mwl} \cdot f_t \cdot UC_{mwt}^{\omega} \\ + \sum_{\omega mwlt} p_m^{\omega} \cdot wg_{mwl} \cdot v_t \cdot P_{mwlt}^{\omega} \\ + \sum_{\omega mwl} p_m^{\omega} \cdot wg_{mwl} \cdot v' \cdot ENS_{mwl}^{\omega} \\ + \sum_{\omega mwl} p_m^{\omega} \cdot wg_{mwl} \cdot v'' \cdot RNS_{mwl}^{\omega}\end{aligned} \tag{B.3-1}$$

Subject to:

$$\sum_g O_{mwlg}^{\omega} + RNS_{mwl}^{\omega} \geq o_{mwl} \quad \forall \omega mwl \tag{B.3-2}$$

$$\sum_g P_{mwlg}^{\omega} - \sum_s \frac{C_{mwls}^{\omega}}{\eta_s} + ENS_{mwl}^{\omega} = d_{mwl} \quad \forall \omega mwl \tag{B.3-3}$$

$$P_{mwl+1g}^{\omega} \leq P_{mwlg}^{\omega} \quad \forall \omega mwlg \tag{B.3-4}$$

$$UC_{mwt}^{\omega} - UC_{m-1w+1t}^{\omega'} = SU_{mwt}^{\omega} - SD_{mwt}^{\omega} \quad \forall \omega mwt \quad \omega' \epsilon\, a(\omega) \tag{B.3-5}$$

$$UC_{mw+1t}^{\omega} - UC_{mwt}^{\omega} = SU_{mw+1t}^{\omega} - SD_{mw+1t}^{\omega} \quad \forall \omega mt\ \forall w > 1 \tag{B.3-6}$$

$$P_{mwlt}^{\omega} + O_{mwlt}^{\omega} \leq \bar{p}_t UC_{mwt}^{\omega} \quad \forall \omega mwt \tag{B.3-7}$$

$$P_{mwlt}^{\omega} \geq \underline{p}_t UC_{mwt}^{\omega} \quad \forall \omega mwt \tag{B.3-8}$$

$$P_{mwlh}^{\omega} \leq \bar{p}_h \quad \forall \omega mwh \tag{B.3-9}$$

$$\begin{aligned}R_{m-1r}^{inter,\omega'} - R_{mr}^{inter,\omega} + i_{mr}^{\omega} - S_{mr}^{\omega} + \underbrace{\sum_{r' \in RUR_{r',r}} S_{mr'}^{\omega}}_{Water\ spillage\ from\ upstream\ reservoirs} \\ + \underbrace{\sum_{wl} \sum_{h \in HUR_{h,r}} wg_{mwl} \cdot P_{mwlh}^{\omega} / c_h}_{Turbined\ water\ from\ upstream\ hydro\ plants} \\ - \underbrace{\sum_{wl} \sum_{h \in RUH_{r,h}} wg_{mwl} \cdot P_{mwlh}^{\omega} / c_h}_{Turbined\ water\ from\ hydro\ plants\ in\ reservoir\ r} \\ + \underbrace{\sum_{wl} \sum_{h \in HPR_{h,r}} wg_{mwl} \cdot C_{mwlh}^{\omega} / c_h}_{Pumped\ water\ from\ hydro\ plants\ to\ reservoir\ r} \\ - \underbrace{\sum_{wl} \sum_{h \in RPH_{r,h}} wg_{mwl} \cdot C_{mwlh}^{\omega} / c_h}_{Pumped\ water\ to\ other\ reservoirs} = 0 \quad \forall \omega mr \quad \omega' \epsilon\, a(\omega)\end{aligned} \tag{B.3-10}$$

$$SoC_{m-1b}^{inter,\omega'} - SoC_{mb}^{inter,\omega} - \sum_{wl} wg_{mwl} \cdot P_{mwlb}^{\omega} + \sum_{wl} wg_{mwl} \cdot C_{mwlb}^{\omega} = 0 \quad \forall \omega mb \quad \omega' \epsilon\, a(\omega) \tag{B.3-11}$$

$$R_{|M|r}^{inter,\omega} = r_r' \quad \forall \omega r \tag{B.3-12}$$

$$\underline{r}_r \leq R_{mr}^{inter,\omega} \leq \overline{r}_r \quad \forall \omega mr \tag{B.3-13}$$

$$\underline{soc}_b \leq SoC_{mb}^{inter,\omega} \leq \overline{soc}_b \quad \forall \omega mb \tag{B.3-14}$$

$$0 \leq P_{mwlg}^{\omega} \leq \overline{p}_g \quad \forall \omega mwg \tag{B.3-15}$$

$$0 \leq C_{mwls}^{\omega} \leq \overline{p}_s \quad \forall \omega mws \tag{B.3-16}$$

$$UC_{mwt}^{\omega}, SU_{mwt}^{\omega}, SU_{mwt}^{\omega} \in \{0,1\} \quad \forall \omega mwt \tag{B.3-17}$$

The objective function (B.3-1) minimizes the expected operational costs such as (B.2-1) in the HM model. Constraints (B.3-2) to (B.3-17) have the same purpose as in the HM model. However, they are stated in term of the LDC time division, i.e., m, w, l. Moreover, the following considerations are particular for the LDC model: Equation (B.3-4) limits the production in consecutive load levels. Equations (B.3-5) and (B.3-6) represent the commitment constraints in the LDC model. This model considers startup and shutdown decisions within aggregation of load levels w. For instance, if the aggregations of load levels are weekdays and weekend, then equations (B.3-5) and (B.3-6) define the commitment, startup, and shutdown decisions between weekdays and the following weekend and vice versa.

Equation (B.3-10) defines the water balance for each reservoir considering the hydro topology while (B.3-11) defines the state of charge for each short-term storage (e.g., batteries). Both equations include the load-level duration (wg_{mwl}) to consider the number of hours that are represented for each load level. In other words, the multiplication by wg_{mwl} guarantees that all the charged/discharged energy is considered within the month m. These equations are for the inter-period variables. Intra-period variables are not available in this model due to the lack of chronology within the month m.

B.4 LRP Stochastic Hydrothermal Dispatch

This section describes the LRP model, which is a stochastic extension of the previous formulation developed in [95]. This model is a novel contribution to hydrothermal dispatch because it enables consideration of both short-term storage (e.g., batteries) and seasonal storage simultaneously.

$$\min \sum_{\omega m} p_m^{\omega} \sum_{rp,k,t} wg_{rp} \cdot su_t \cdot SU_{rp,k,t}^{\omega} + \sum_{\omega m} p_m^{\omega} \sum_{t,(rp,k)\in\{CI_{p,rp,k}\cap MP_{m,p}\}} wg_{rp} \cdot sd_t \cdot SD_{rp,k,t}^{\omega} + \sum_{\omega m} p_m^{\omega} \sum_{t,(rp,k)\in\{CI_{p,rp,k}\cap MP_{m,p}\}} wg_{rp} \cdot f_t \cdot UC_{rp,k,t}^{\omega} + \sum_{\omega m} p_m^{\omega} \sum_{t,(rp,k)\in\{CI_{p,rp,k}\cap MP_{m,p}\}} wg_{rp} \cdot v_t \cdot P_{rp,k,t}^{\omega} + \sum_{\omega m} p_m^{\omega} \sum_{(rp,k)\in\{CI_{p,rp,k}\cap MP_{m,p}\}} wg_{rp} \cdot v' \cdot ENS_{rp,k}^{\omega} + \sum_{\omega m} p_m^{\omega} \sum_{(rp,k)\in\{CI_{p,rp,k}\cap MP_{m,p}\}} wg_{rp} \cdot v'' \cdot RNS_{rp,k}^{\omega} \quad \text{(B.4-1)}$$

Subject to:

$$\sum_g O_{rp,k,g}^{\omega} + RNS_{rp,k}^{\omega} \geq o_{rp,k} \quad \forall \omega, rp, k \quad \text{(B.4-2)}$$

$$\sum_g P_{rp,k,g}^{\omega} - \sum_s \frac{C_{rp,k,s}^{\omega}}{\eta_s} + ENS_{rp,k}^{\omega} = d_{rp,k} \quad \forall \omega, rp, k \quad \text{(B.4-3)}$$

$$UC_{rp,k,t}^{\omega} - UC_{rp,k-1,t}^{\omega} = SU_{rp,k,t}^{\omega} - SD_{rp,k,t}^{\omega} \quad \forall \omega, rp, t \; \forall k > 1 \quad \text{(B.4-4)}$$

$$UC_{rp,k,t}^{\omega} - \sum_{rp'\in TM_{rp',rp}} UC_{rp',K,t}^{\omega'} = SU_{rp,k,t}^{\omega} - SD_{rp,k,t}^{\omega} \quad \forall \omega, rp, t \quad \forall k = 1 \quad \omega' \epsilon \, a(\omega) \quad \text{(B.4-5)}$$

$$P'^{\omega}_{rp,k,t} + O_{rp,k,t}^{\omega} \leq \left(\bar{p}_t - \underline{p}_t\right) \cdot \left(UC_{rp,k,t}^{\omega} - SU_{rp,k,t}^{\omega}\right) \quad \forall \omega, rp, k, t \quad \text{(B.4-6)}$$

$$P'^{\omega}_{rp,k,t} + O_{rp,k,t}^{\omega} \leq \left(\bar{p}_t - \underline{p}_t\right) \cdot \left(UC_{rp,k,t}^{\omega} - SD_{rp,k+1,t}^{\omega}\right) \quad \forall \omega, rp, k, t \quad \text{(B.4-7)}$$

$$P_{rp,k,t}^{\omega} = \underline{p}_t UC_{rp,k,t}^{\omega} + P'^{\omega}_{rp,k,t} \quad \forall \omega, rp, k, t \quad \text{(B.4-8)}$$

$$P_{rp,k,h}^{\omega} \leq \bar{p}_h \quad \forall \omega, rp, k, h \quad \text{(B.4-9)}$$

$$
\begin{aligned}
&R^{intra,\omega'}_{rp,k-1,r} - R^{intra,\omega}_{rp,k,r} + i^{\omega}_{rp,k,r} - S^{\omega}_{rp,k,r} \\
&\quad + \underbrace{\sum_{r' \in RUR_{r',r}} S^{\omega}_{rp,k,r'}}_{Water\ spillage\ from\ upstream\ reservoirs} \\
&\quad + \underbrace{\sum_{h \in HUR_{h,r}} P^{\omega}_{rp,k,h}/_{c_h}}_{Turbined\ water\ from\ upstream\ hydro\ plants} \\
&\quad - \underbrace{\sum_{h \in RUH_{r,h}} P^{\omega}_{rp,k,h}/_{c_h}}_{Turbined\ \ water\ from\ hydro\ plants\ in\ reservoir\ r} \\
&\quad + \underbrace{\sum_{h \in HPR_{h,r}} C^{\omega}_{rp,k,h}/_{c_h}}_{Pumped\ water\ from\ hydro\ plants\ to\ reservoir\ r} \\
&\quad - \underbrace{\sum_{h \in RPH_{r,h}} C^{\omega}_{rp,k,h}/_{c_h}}_{Pumped\ water\ to\ other\ reservoirs} \\
&\quad = 0 \quad \forall \omega, rp, k, r \quad \omega' \epsilon\ a(\omega)
\end{aligned}
\tag{B.4-10}
$$

$$
\begin{aligned}
&R^{inter,\omega'}_{m-1,r} - R^{inter,\omega}_{mr} + \sum_{(rp,k) \in \{CI_{p,rp,k} \cap MP_{m,p}\}} \left[R^{intra,\omega}_{rp,k,r} - R^{intra,\omega'}_{rp,k-1,r}\right] \\
&\quad = 0 \quad \forall \omega mr \quad \omega' \epsilon\ a(\omega)
\end{aligned}
\tag{B.4-11}
$$

$$
SoC^{intra,\omega'}_{rp,k-1,b} - SoC^{intra,\omega}_{rp,k,b} - P^{\omega}_{rp,k,b} + C^{\omega}_{rp,k,b} = 0 \quad \forall \omega, rp, k, b \quad \omega' \epsilon\ a(\omega) \tag{B.4-12}
$$

$$
\begin{aligned}
&SoC^{inter,\omega'}_{m-1,b} - SoC^{inter,\omega}_{mb} \\
&\quad + \sum_{(rp,k) \in \{CI_{p,rp,k} \cap MP_{m,p}\}} \left[SoC^{intra,\omega}_{rp,k,r} - SoC^{intra,\omega'}_{rp,k-1,r}\right] \\
&\quad = 0 \quad \forall \omega mb \quad \omega' \epsilon\ a(\omega)
\end{aligned}
\tag{B.4-13}
$$

$$
R^{inter,\omega}_{|M|r} = r'_r \quad \forall \omega r \tag{B.4-14}
$$

$$
\underline{r}_r \leq R^{inter,\omega}_{mr} \leq \overline{r}_r \quad \forall \omega mr \tag{B.4-15}
$$

$$
\underline{r}_r \leq R^{intra,\omega}_{rp,k,r} \leq \overline{r}_r \quad \forall \omega, rp, k, r \tag{B.4-16}
$$

$$
\underline{soc}_b \leq SoC^{inter,\omega}_{mb} \leq \overline{soc}_b \quad \forall \omega mb \tag{B.4-17}
$$

$$
\underline{soc}_b \leq SoC^{intra,\omega}_{rp,k,b} \leq \overline{soc}_b \quad \forall \omega, rp, k, b \tag{B.4-18}
$$

$$
0 \leq P^{\omega}_{rp,k,g} \leq \overline{p}_g \quad \forall \omega, rp, k, g \tag{B.4-19}
$$

$$
0 \leq C^{\omega}_{rp,k,s} \leq \overline{p}_s \quad \forall \omega, rp, k, s \tag{B.4-20}
$$

$$
UC^{\omega}_{rp,k,t}, SU^{\omega}_{rp,k,t}, SU^{\omega}_{rp,k,t} \in \{0,1\} \quad \forall \omega, rp, k, t \tag{B.4-21}
$$

The objective function (B.4-1) also minimizes the expected operational costs such as (B.2-1) in the HM model. The operational costs associated with each RP are multiplied by the number of periods in the time horizon that are represented by it, i.e., multiplied by the weight of each RP. In addition, the intersection of both sets $\{CI_{p,rp,k} \cap MP_{m,p}\}$ guarantees that we are considering only the RPs belonging to the corresponding month. Constraints (B.4-2) to

(B.4-21) have the same purpose as their equivalent ones in the HM model. However, they are stated in term of the RPs time division, i.e., rp, k. Moreover, the following considerations are specific to the LRD model: Equations (B.4-4) and (B.4-5) are the commitment constraints in the LRD model. Equation (B.4-4) is for all the hours inside the RP except for the first hour. For the first hour, (B.4-5) creates continuity between the RPs and prevents unnecessary startups by using a transition matrix, i.e., $TM_{rp',rp}$, to require that for any pair of RPs that transition from one to the other, the thermal unit status in the last hour of the first, i.e., rp', is considered in the first hour of the second, i.e., rp.

Equations (B.4-10) to (B.4-13) represent the balance equations for both types of storage, i.e., hydro reservoirs and batteries. These equations create the continuity in storage across the entire time horizon that allows for the modeling of short-term and long-term storage simultaneously. Equations (B.4-10) and (B.4-12) ensure the storage balance within the RP, while (B.4-11) and (B.4-13) guarantee the storage balance by checking at regular intervals (e.g., aggregation of hours such as months m) that all the energy charged and discharged since the previous month plus the total energy at the last checkpoint are within bounds. This is possible because the cluster index, $CI_{p,rp,k}$, and the relationship between periods and months, $MP_{m,p}$, are known as a result of the procedure to determine the RPs. The intersection of both sets $\{CI_{p,rp,k} \cap MP_{m,p}\}$ indicates which RPs belong to the month and, therefore, must be considered in the inter-period balance.

C ESS ALLOCATION AND INVESTMENT FORMULATIONS

C.1 Notation

Indices and Sets

$n, m \in N$	Buses
$j \in J$	Storage technologies
$t \in T$	Time intervals
$k \in K$	Linear approximation segments

Parameters

η_j^c	Charging efficiency of technologies [p.u.]
η_j^d	Discharging efficiency of technologies [p.u]
Δt	Duration of time step [h]
R_j^c	Maximum charging capacity [MW]
R_j^d	Maximum discharging capacity [MW]
P_n^{min}	Minimum thermal power output [MW]
P_n^{max}	Maximum thermal power output [MW]
$+RR_n;\ -RR_n$	Upper and lower ramp rate limits [MW]
$W_{n(t)}$	Total wind generation [MW]
TC_{nm}^{max}	Maximum line transmission capacity [MW]
Gt_{nm}	Circuit Conductance [p.u.]
$C_n^{g1}(t)$	Linear production cost coefficient [$/MWh]
C_n^{g2}	Quadratic production cost coefficient [$/MWh^2]
$C_{jn}^{\mathrm{d1}}(t)$	Linear discharging cost coefficient [$/MWh]
C_{jn}^{d2}	Quadratic discharging cost coefficient [$/MWh^2]
C_j^i	Investment cost [$/MW]
DD_j	Average discharge duration of each technology [h]
$D_n(t)$	Demand [MW]
Y_k	Quadratic function values for linearization
X_k	Original x-axis values used for linearization
I_k	y-intercept values for each segment in linearization

Variables

$p_n^g(t)$	Generation from thermal units [MW]
$r_{jn}^c(t)$	Charging rates of each storage technology [MW]
$r_{jn}^d(t)$	Discharging rates of each storage technology [MW]
$s_{jn}(t)$	Storage level [MWh]
k_{jn}	Installed storage capacity [MWh]
$\delta_n(t)$	Voltage angle at a node [p.u.]
$\Delta_{nm}(t)$	Difference in voltage angles between two buses [rad]
$\Delta'_{nm}(t)$	Square of difference in voltage angles between two buses

C.2 Storage Allocation Model

The decision variables used throughout the DC OPF storage optimization model comprise the set:

$$\Omega := \left\{ p_n^g(t), r_{jn}^c(t), r_{jn}^d(t), s_{jn}(t), k_{jn}, \delta_n(t), \Delta_{nm}(t), \Delta'_{nm}(t) \right\}$$

The objective function of the model calculates the total system costs:

$$\min_{\Omega} \sum_{t \in T} \left\{ \sum_{n \in G} F_n^g\left(p_n^g(t), t\right) + \sum_{n \in N, j \in J} F_{jn}^d\left(r_{jn}^d(t), t\right) \right\} \tag{C.2-1}$$

This objective is comprised of the production cost function:

$$F_n^g\left(p_n^g(t), t\right) := C_n^{g1}(t) p_n^g(t) \Delta t + C_n^{g2}(t) \left(p_n^g(t) \Delta t\right)^2$$

and the discharging costs:

$$F_{jn}^d\left(r_{jn}^d(t), t\right) := C_{jn}^{d1}(t) r_{jn}^d(t) \Delta t + C_{jn}^{d2}(t) \left(r_{jn}^d(t) \Delta t\right)^2$$

which together represent the fixed and variable costs associated with network storage. For thermal generation units, convex quadratic cost functions are commonly used in unit commitment formulations [133]. While this study does not take unit commitment into consideration, the solution of such a formulation results in a model that is straightforward using conic optimization algorithms. For this reason, it was adopted here. Cost functions associated with charging and discharged stored power have been included by other authors in previous studies [122], [152]. These allow the model to take into account optimal system deployment as well as allocation and investment.

Let us discuss the constraints used in formulating this optimization problem. The storage level in each technology will be defined for $\forall j, n, t \geq 2$ such that

$$s_{jn}(t) = s_{jn}(t-1) + \left(\eta_j^c r_{jn}^c(t) - r_{jn}^d(t)/\eta_j^d\right) \Delta t \tag{C.2-2}$$

This states that the energy storage level of each technology at each node during every time step past the second one, must be equal to the storage in the previous period plus the difference of the energy charged and discharged between the previous period and the current time step. Additionally, the problem variables were bounded as follows:

$$0 \leq r_{jn}^c(t) \leq R_j^c \qquad \forall j, n, t \qquad \text{(C.2-3)}$$

$$0 \leq r_{jn}^d(t) \leq R_j^d \qquad \forall j, n, t \qquad \text{(C.2-4)}$$

$$0 \leq s_{jn}(t) \leq k_{jn} \qquad \forall j, n, t \qquad \text{(C.2-5)}$$

$$s_{jn}(t=1) = s_{jn}(t=T) \qquad \forall j, n \qquad \text{(C.2-6)}$$

$$P_n^{min} \leq p_n^g(t) \leq P_n^{max} \qquad \forall j, n \qquad \text{(C.2-7)}$$

$$-RR_n \leq p_n^g(t) - p_n^g(t-1) \leq +RR_n \qquad \forall j, n \qquad \text{(C.2-8)}$$

$$k_{jn} = 0 \qquad \forall j, n \in GR_{jn} \qquad \text{(C.2-9)}$$

$$\sum_{n \in N} k_{jn} \leq SC_j^{max} \qquad \forall j \qquad \text{(C.2-10)}$$

$$\Delta_{nm}(t) = \delta_n(t) - \delta_m(t) \qquad \forall m \in \Theta_n, t \qquad \text{(C.2-11)}$$

Bounds (C.2-3) and (C.2-4) define the upper and lower limits for charging and discharging respectively. Constraint (C.2-5) states that the storage level for a technology cannot exceed the amount of installed capacity of technology j at node n during time t. Constraint (C.2-6) states the ending energy storage level of a technology must be the same as the starting value. Bound (C.2-7) defines the minimum and maximum thermal generation limits. (C.2-8) defines the upper and lower ramping limitations of thermal generators. Let GR_{jn} be defined as the set of nodes where geographical, zoning or social issues prevent the installation of a particular technology at that node. For example, a pumped hydro storage facility cannot work in an area without abundant water. Consequently, (C.2-9) states that no capacity can be installed for a technology at a node that falls within that set. Bound (C.2-10) states that the total installed capacity over all nodes for technology j cannot exceed SC_j^{max}, which denotes the total available storage capacity of a given technology. Lastly (C.2-11) defines the voltage angle difference for the set Θ_n, of buses m connected to bus n.

The energy at each node n during each time period t can be defined as an equality between the inflows and outflows of power. This model assumes that all generated wind was consumed, and therefore no wind curtailment was observed. Therefore, wind could be treated as input data. Transmission losses are given by the exact active power dissipated, $2Gt_{nm}\big(1-\cos\big(\Delta_{nm}(t)\big)\big)$ and are linearized as in Section 5.1.2.

$$W_{n(t)} + p_n^g(t) + \sum_{j\in J} r_{jn}^d(t) = D_n(t) + \sum_{m\in\Theta_n} B_{nm}\big(\delta_n(t)-\delta_m(t)\big) + \sum_{j\in J} r_{jn}^c(t) + 2Gt_{nm}\big(1-\cos\big(\Delta_{nm}(t)\big)\big) \qquad \forall\, n,t \quad \text{(C.2-12)}$$

Transmission limits needed to be put in place to account for network congestion. (C.2-13) limits the power flow between two nodes to the line limit as TC_{nm}^{max}. The set Θ_n defined previously will be used here to ensure that only lines that exist are bounded.

$$-TC_{nm}^{max} \le B_{nm}\big(\delta_n(t)-\delta_m(t)\big) \le TC_{nm}^{max} \qquad \forall\, m \in \Theta_n, t \quad \text{(C.2-13)}$$

The corresponding voltage angles are bounded by:

$$-\pi \le \delta_n(t) \le \pi \qquad \forall\, n,t \quad \text{(C.2-14)}$$

Finally, $n = 1$ is defined as the slack bus for all time steps t by:

$$\delta_{n=1}(t) = 0 \qquad \forall\, t \quad \text{(C.2-15)}$$

The model given by (C.2-1) to (C.2-15) represents an optimization problem that is to be solved in the storage allocation problem. However, this model is non-convex and non-linear and will be adjusted in the following section.

C.3 Transmission Losses Linearization

To maintain the optimization model's tractability, a linearized piecewise ohmic loss approximation is used. Section 5.1.2 have shown the detailed explanation of the linearization method. Here, we summarized the main equations. Equation (C.3-1) allows us to calculate the equations of each segment k used to approximate the quadratic function with linear curves at specific points.

$$\Delta'_{nm}(t) \geq \left(2X_k \Delta_{nm}(t)\right) + I_k \qquad \forall k, m \in \Theta_n, t \qquad \text{(C.3-1)}$$

Using the linear segments obtained with (C.3-1), we find that the values of $\Delta'_{nm}(t)$ are approximately equal to the square of the voltage angle differences used in the general quadratic approximation. Transmission losses on a line can then be represented as $0.5\left(Gt_{nm}\Delta'_{nm}(t)\right)$. Note that product is halved as to prevent counting both the positive and negative sides of the approximated curve. Therefore, we must adjust (C.2-12) to represent the linearized approximation of losses in the system energy balance. The new equation for this is displayed in (C.3-2).

$$\begin{aligned} W_{n(t)} + p_n^g(t) + \sum_{j \in J} r_{jn}^d(t) \\ = D_n(t) + \sum_{m \in \Theta_n} B_{nm}\left(\delta_n(t) - \delta_m(t)\right) \\ + \sum_{j \in J} r_{jn}^c(t) + \sum_{m \in \Theta_n} 0.5\left(Gt_{nm}\Delta'_{nm}(t)\right) \end{aligned} \qquad \forall\, n, t \qquad \text{(C.3-2)}$$

The linear optimization problem to solve the storage allocation, includes (C.2-1) to (C.2-15), but replaces (C.2-12) with (C.3-2) and incorporates additional constraint (C.3-1).

C.4 Storage Investment Model

The model can now be extended to consider the investment costs of new storage capacity. Until now, the available capacity was assumed fixed for each technology, therefore implicitly accounting for an investment decision. The new objective function will include the cost function of storage investment:

$$F_j^i\left(k_{jn}\right) = k_{jn}^i C_j^i / DD_j$$

The objective function is now given by (C.4-1) and subject to constraints (C.2-2)–(C.2-11), (C.2-13)–(C.2-15), (C.3-1)–(C.3-2). This model portrays a convex optimization problem with linear constraints, and a quadratic objective function. In addition, k_{jn} are treated as continuous values.

$$\min_{\Omega} \sum_{t \in T} \left\{ \sum_{n \in G} F_n^g\left(p_n^g(t), t\right) + \sum_{n \in N, j \in J} F_{jn}^d\left(r_{jn}^d(t), t\right) \right\} + \sum_{n \in N, j \in J} F_j^i\left(k_{jn}\right) \qquad \text{(C.4-1)}$$

D POWER- AND ENERGY-BASED FORMULATIONS

D.1 Notation

Indices and Sets

$j \in \mathcal{J}$	Technologies
$g \in \mathcal{G} \subseteq \mathcal{J}$	Subset of thermal generation technologies
$v \in \mathcal{V} \subseteq \mathcal{J}$	Subset of renewable energy sources
$s \in \mathcal{S} \subseteq \mathcal{J}$	Subset of energy storage technologies
$b \in \mathcal{B}$	Buses
$\mathcal{B}^D \subseteq \mathcal{B}$	Subset of buses b with demand consumption
$l \in \mathcal{L}$	Transmission lines
$\omega \in \Omega$	Scenarios
$k \in \mathcal{K}_g$	Startup segments, running from 1 (the hottest) to K_g (the coldest)
$t \in \mathcal{T}$	time periods (e.g., hours)

Parameters

C_j^{LV}	Linear variable production cost [$/MWh]
C_g^{NL}	No-load cost [$/h]
C_g^{SD}	Shutdown cost [$]
C_{gk}^{SU}	Startup cost for segment k [$]
C_g^{EM}	CO2 emission cost [$/MWh]
C_j^{R+}, C_j^{R-}	Up/down reserve cost [$/MW]
$D_{\omega bt}^E$	Energy demand on bus b [MWh]
$D_{\omega bt}^P$	Power demand on bus b [MW]
$R_{\omega t}^+, R_{\omega t}^-$	Up/down reserve requirement [MW]
$\overline{F}_l$	Power flow limit on transmission line l [MW]
$\overline{P}_g, \underline{P}_g$	Maximum/minimum power output [MW]
$E_{gkt}^{SU}, E_{gt}^{SD}$	Energy output during startup/shutdown [MWh]
$P_{gkt}^{SU}, P_{gt}^{SD}$	Power output during startup/shutdown [MW]
RU_g, RD_g	Ramp-up/down capability [MW/min]
SU_g, SD_g	Startup/shutdown capability [MW]
SU_g^D, SD_g^D	Startup/shutdown duration [h]
T_{gk}^{SU}	Time interval limit of startup segment k [h]
TU_g, TD_g	Minimum up/down time [h]
$\Gamma_{lj}^J, \Gamma_{lb}$	Shift factors for line l [p.u.]
EPR_s	Energy to power ratio [h]
$V_{\omega vt}^E$	Renewable energy output profile [p.u.]
$V_{\omega vt}^P$	Renewable power output profile [p.u.]
π_ω	Probability of scenario ω
$\overline{X}_j$	Investment limit for technology j
X_j^0	Initial capacity for technology j; [# units] for g, and [MW] for s and v.

Continuous Non-negative Variables

$\hat{e}_{\omega jt}$	Total energy output [MWh]
$\hat{p}_{\omega jt}$	Total power output [MW]
$e_{\omega gt}$	Energy output above minimum output [MWh]
$p_{\omega gt}$	Power output above minimum output [MW]
$\hat{c}_{\omega st}$	Charged energy for storage [MWh]
$c_{\omega st}$	Charged power for storage [MW]
$r^{+}_{\omega gt}$	Up capacity reserve [MW]
$r^{-}_{\omega gt}$	Down capacity reserve [MW]
$\phi_{\omega st}$	Energy storage level [MWh]

Integer Variables

$u_{\omega gt}$	Unit commitment for thermal technologies
$y_{\omega gt}$	Startup for thermal technologies
$z_{\omega gt}$	Shutdown for thermal technologies
$\delta_{\omega gkt}$	Startup type selection for thermal technologies
$\gamma_{\omega st}$	Binary decision for charging/discharging logic
x_j	Investment decision per technology

D.2 Energy-Based GEP-UC Model

The GEP seeks to minimize the investment costs plus the expected value of operating costs: production cost, up/down reserve cost, CO2 emission cost, no-load cost, shutdown cost, startup cost. Notice that $\Psi = \{x, e, \hat{e}, \hat{c}, r^{+}, r^{-}, u, y, z, \delta, \phi\}$ corresponds to the set of decision variables considered in this model.

$$\min_{\Psi} \sum_{j\in\mathcal{J}} C_j^I x_j + \sum_{\omega\in\Omega} \pi_\omega \sum_{t\in\mathcal{T}} \left\{ \sum_{j\in\mathcal{J}} \left[C_j^{LV} \hat{e}_{\omega jt} + C_j^{R+} r^{+}_{\omega jt} + C_j^{R-} r^{-}_{\omega jt} \right] + \sum_{g\in\mathcal{G}} \left[C_g^{EM} \hat{e}_{\omega gt} + C_g^{NL} u_{\omega gt} + C_g^{SD} z_{\omega gt} + \sum_{k\in\mathcal{K}_g} C_{gk}^{SU} \delta_{\omega gkt} \right] \right\} \quad \text{(D.2-1)}$$

The system-wide constraints are guaranteed by energy demand balance (D.2-2), transmission limits (D.2-3), and reserve requirements (D.2-4)-(D.2-5):

$$\sum_{j\in\mathcal{J}} \hat{e}_{\omega jt} - \sum_{s\in\mathcal{S}} \hat{c}_{\omega st} = \sum_{b\in\mathcal{B}^D} D^E_{\omega bt} \quad \forall \omega, t \quad \text{(D.2-2)}$$

$$-\overline{F}_l \le \sum_{j\in\mathcal{J}} \Gamma^J_{lj} \hat{e}_{\omega jt} - \sum_{s\in\mathcal{S}} \Gamma^S_{ls} c_{\omega st} - \sum_{b\in\mathcal{B}^D} \Gamma_{lb} D^E_{\omega bt} \le \overline{F}_l \quad \forall l, \omega, t \quad \text{(D.2-3)}$$

$$\sum_{j\in\mathcal{J}} r^{+}_{\omega jt} \geq R^{+}_{\omega t} \quad \forall \omega, t \tag{D.2-4}$$

$$\sum_{j\in\mathcal{J}} r^{-}_{\omega jt} \geq R^{-}_{\omega t} \quad \forall \omega, t \tag{D.2-5}$$

The relationship between operational and investment decisions for each technology type is guaranteed with (D.2-6) for thermal technologies, (D.2-7)-(D.2-8) for ESS, and (D.2-9) for VRES.

$$u_{\omega gt} \leq X^{0}_{g} + x_{g} \quad \forall \omega, g, t \tag{D.2-6}$$

$$\hat{e}_{\omega st} - \hat{c}_{\omega st} + r^{+}_{\omega gt} \leq X^{0}_{s} + x_{s} \quad \forall \omega, s, t \tag{D.2-7}$$

$$\hat{e}_{\omega st} - \hat{c}_{\omega st} - r^{-}_{\omega gt} \geq -(X^{0}_{s} + x_{s}) \quad \forall \omega, s, t \tag{D.2-8}$$

$$\hat{e}_{\omega vt} \leq V^{E}_{\omega vt}(X^{0}_{v} + x_{v}) \quad \forall \omega, v, t \tag{D.2-9}$$

Thermal generation constraints include: commitment/ startup/ shutdown logic (D.2-10), minimum up/down times (D.2-11)-(D.2-12), startup type selection (D.2-13)-(D.2-14) (e.g., hot, warm, and cold startup), energy production limits including reserve decisions (D.2-15)-(D.2-18) (where $\mathcal{G}^{1}$ is defined as the thermal technologies in $\mathcal{G}$ with $TU_g = 1$), and total energy production (D.2-19). The UC formulation presented here is based on the tight and compact formulation proposed in [145]. Furthermore, Gentile et al. [144] have proven that the set of constraints (D.2-10)-(D.2-12) together with (D.2-15)-(D.2-19) is the tightest representation (i.e., convex hull) for the energy-based model.

$$u_{\omega gt} - u_{\omega g,t-1} = y_{\omega gt} - z_{\omega gt} \quad \forall \omega, g, t \tag{D.2-10}$$

$$\sum_{i=t-TU_g+1}^{t} y_{\omega gi} \leq u_{\omega gt} \quad \forall \omega, g, t \in [TU_g, T] \tag{D.2-11}$$

$$\sum_{i=t-TD_g+1}^{t} z_{\omega gi} \leq \left(X^{0}_{g} + x_{g}\right) - u_{\omega gt} \quad \forall \omega, g, t \in [TD_g, T] \tag{D.2-12}$$

$$\delta_{\omega gkt} \leq \sum_{i=T^{SU}_{gk}}^{T^{SU}_{g,k+1}-1} z_{\omega g,t-i} \quad \forall \omega, g, k \in [1, K_g), t \tag{D.2-13}$$

$$\sum_{k\in\mathcal{K}_g} \delta_{\omega gkt} = y_{\omega gt} \quad \forall \omega, g, t \tag{D.2-14}$$

$$e_{\omega gt} + r^{+}_{\omega gt} \leq \left(\overline{P}_g - \underline{P}_g\right)u_{\omega gt} - \left(\overline{P}_g - SD_g\right)z_{\omega g,t+1} - \max\left(SD_g - SU_g, 0\right)y_{\omega gt} \quad \forall \omega, g \in \mathcal{G}^{1}, t \tag{D.2-15}$$

$$e_{\omega gt} + r^{+}_{\omega gt} \leq \left(\overline{P}_g - \underline{P}_g\right)u_{\omega gt} - \left(\overline{P}_g - SU_g\right)y_{\omega g,t} - \max\left(SU_g - SD_g, 0\right)z_{\omega g,t+1} \quad \forall \omega, g \in \mathcal{G}^{1}, t \tag{D.2-16}$$

$$e_{\omega gt} + r^{+}_{\omega gt} \leq \left(\overline{P}_g - \underline{P}_g\right)u_{\omega gt} - \left(\overline{P}_g - SU_g\right)y_{\omega g,t} - \left(\overline{P}_g - SD_g\right)z_{\omega g,t+1} \quad \forall \omega, g \notin \mathcal{G}^{1}, t \tag{D.2-17}$$

$$e_{\omega gt} - r^{-}_{\omega gt} \geq 0 \quad \forall \omega, g, t \tag{D.2-18}$$

$$\hat{e}_{\omega gt} = \underline{P_g} u_{\omega gt} + e_{\omega gt} \quad \forall \omega, g, t \tag{D.2-19}$$

Traditional energy-based UC formulations ignore the inherent startup (SU) and shutdown (SD) trajectories of thermal generation, assuming they start/end their production at their minimum output. Authors in [31], [70] have shown the relevance of the SU and SD processes when they are included in the scheduling optimization. Therefore, we also analyze the energy-based formulation including the SU/SD trajectories proposed in [143]. Thus, if SU/SD trajectories are considered then (D.2-19) is replaced by (D.2-20).

$$\begin{aligned}\hat{e}_{\omega gt} = & \underbrace{\sum_{k=1}^{K_g} \sum_{i=1}^{SU_{gk}^D} E_{gki}^{SU} \delta_{\omega gk,\left(t-i+SU_{gk}^D+1\right)}}_{Startup\ trajectory} + \underbrace{\sum_{i=1}^{SD_g^D} E_{gi}^{SD} z_{\omega g,(t-i+1)}}_{Shutdown\ trajectory} \\ & + \underbrace{\underline{P_g} u_{\omega gt} + e_{\omega gt}}_{Output\ when\ being\ up} \quad \forall \omega, g, t\end{aligned} \tag{D.2-20}$$

ESS constraints include: logic to avoid charging and discharging at the same time (D.2-21)-(D.2-22), the definition of the storage inventory level (D.2-23), storage limits including reserve (D.2-24)-(D.2-25).

$$\hat{c}_{\omega st} \leq (1 - \gamma_{\omega st}) \cdot \left(X_s^0 + \overline{X}_s\right) \quad \forall \omega, s, t \tag{D.2-21}$$

$$\hat{e}_{\omega st} \leq \gamma_{\omega st} \cdot \left(X_s^0 + \overline{X}_s\right) \quad \forall \omega, s, t \tag{D.2-22}$$

$$\phi_{\omega st} = \phi_{\omega s,t-1} + \eta_s \hat{c}_{\omega st} - \hat{e}_{\omega st} \quad \forall \omega, s, t \tag{D.2-23}$$

$$\phi_{\omega st} \leq EPR_s (X_s^0 + x_s) - \textstyle\sum_{i=t-1}^{t} r_{\omega gi}^- \quad \forall \omega, s, t \tag{D.2-24}$$

$$\phi_{\omega st} \geq \textstyle\sum_{i=t-1}^{t} r_{\omega gi}^+ \quad \forall \omega, s, t \tag{D.2-25}$$

Flexibility requirements in the power system are represented by ramping constraints including reserve decisions. In order to guarantee that scheduled reserves are feasible to provide at τ-min (e.g., τ=5) using the energy-based formulation, it is necessary to consider the ramping capability at τ-min. For instance, ramp capability limits imposed with (D.2-26)-(D.2-27) consider the reserve that thermal technologies can provide at τ-min. ESS ramp capability limits (D.2-28)-(D.2-29) consider the charged energy in addition to the energy output (i.e., discharged energy). Notice that (D.2-28)-(D.2-29) allow ESS to switch from charging to discharging within the ramp limit, as in [142].

$$\left(e_{\omega gt} + r_{\omega gt}^+\right) - e_{\omega g,t-1} \leq \tau RU_g u_{\omega gt} \quad \forall \omega, g, t \tag{D.2-26}$$

$$\left(e_{\omega gt} - r_{\omega gt}^-\right) - e_{\omega g,t-1} \geq -\tau RD_g u_{\omega g,t-1} \quad \forall \omega, g, t \tag{D.2-27}$$

$$\begin{aligned}&\left(\hat{e}_{\omega st} - \hat{e}_{\omega s,t-1}\right) - \left(\hat{c}_{\omega st} - \hat{c}_{\omega s,t-1}\right) + r_{\omega st}^+ \leq \\ &\ \tau RU_g (X_s^0 + x_s) \quad \forall \omega, s, t\end{aligned} \tag{D.2-28}$$

$$\begin{aligned}&\left(\hat{e}_{\omega st} - \hat{e}_{\omega s,t-1}\right) - \left(\hat{c}_{\omega st} - \hat{c}_{\omega s,t-1}\right) - r_{\omega st}^- \geq \\ &\ -\tau RD_g (X_s^0 + x_s) \quad \forall \omega, s, t\end{aligned} \tag{D.2-29}$$

D.3 Power-Based GEP-UC Model

This section shows the GEP-UC equations in terms of power. However, some of the terms in these equations are naturally linked to energy. For instance, the objective function (D.3-1) considers the so-called calculated energy $\hat{e}_{\omega jt}$ to

determine the variable cost and CO2 emission cost. The calculated energy is determined using the power output variables $\hat{p}_{\omega jt}$ in (D.3-2). In addition, for ESS the charged energy $\hat{c}_{\omega st}$ is also determined using the charged power in (D.3-3). Notice that $\Lambda = \{x, p, \hat{p}, \hat{e}, c, \hat{c}, r^+, r^-, u, y, z, \delta, \phi\}$ corresponds to the set of decision variables in this model.

$$\min_{\Lambda} \sum_{j\in\mathcal{J}} C_j^I x_j + \sum_{\omega\in\Omega} \pi_\omega \sum_{t\in\mathcal{T}} \left\{ \sum_{j\in\mathcal{J}} \left[C_j^{LV} \hat{e}_{\omega jt} + C_j^{R+} r_{\omega jt}^+ + C_j^{R-} r_{\omega jt}^- \right] + \sum_{g\in\mathcal{G}} \left[C_g^{EM} \hat{e}_{\omega gt} + C_g^{NL} u_{\omega gt} + C_g^{SD} z_{\omega gt} + \sum_{k\in\mathcal{K}_g} C_{gk}^{SU} \delta_{\omega gkt} \right] \right\} \quad \text{(D.3-1)}$$

$$\hat{e}_{\omega jt} = \frac{\hat{p}_{\omega jt} + \hat{p}_{\omega j,t-1}}{2} \quad \forall \omega, j, t \quad \text{(D.3-2)}$$

$$\hat{c}_{\omega st} = \frac{c_{\omega st} + c_{\omega s,t-1}}{2} \quad \forall \omega, s, t \quad \text{(D.3-3)}$$

Demand balance constraint (D.3-4) and power-flow transmission limits (D.3-5) also use the power output instead of energy output. Reserve requirements (D.2-4)-(D.2-5) remain the same because they are already expressed in terms of power.

$$\sum_{j\in\mathcal{J}} \hat{p}_{\omega jt} - \sum_{s\in\mathcal{S}} c_{\omega st} = \sum_{b\in\mathcal{B}^D} D_{\omega bt}^E \quad \forall \omega, t \quad \text{(D.3-4)}$$

$$-\overline{F}_l \le \sum_{j\in\mathcal{J}} \Gamma_{lj}^J \hat{p}_{\omega jt} - \sum_{s\in\mathcal{S}} \Gamma_{ls}^S c_{\omega st} - \sum_{b\in\mathcal{B}^D} \Gamma_{lb} D_{\omega bt}^E \le \overline{F}_l \quad \forall l, \omega, t \quad \text{(D.3-5)}$$

In terms of the relationship between operational and investment decisions, thermal unit constraint (D.2-6) remains the same. However, constraints for ESS and VRES technologies change to (D.3-6)-(D.3-7) and (D.3-8), respectively.

$$\hat{p}_{\omega st} - c_{\omega st} + r_{\omega gt}^+ \le X_s^0 + x_s \quad \forall \omega, s, t \quad \text{(D.3-6)}$$

$$\hat{p}_{\omega st} - c_{\omega st} - r_{\omega gt}^- \ge -(X_s^0 + x_s) \quad \forall \omega, s, t \quad \text{(D.3-7)}$$

$$\hat{p}_{\omega vt} \le V_{\omega vt}^P (X_v^0 + x_v) \quad \forall \omega, v, t \quad \text{(D.3-8)}$$

Unit commitment constraints (D.2-10)-(D.2-14) do not change in the power-based formulation. Equations (D.3-9)-(D.3-10) limit the power output of thermal technologies. The total power output constraint is different depending whether it is a quick- or slow-start unit. Quick-start technologies $\mathcal{G}^F$ are thermal generators that can startup/shutdown within one hour (i.e., $SU_{gk}^D = SD_g^D \le 1$), while slow-start technologies $\mathcal{G}^S$ are those with a SU/SD duration greater than one hour as well as a SU/SD capacity equal to the minimum power output (i.e., $SU_g = SD_g = \underline{P}_g$). Therefore, the total power output of slow-start technologies considers SU/SD trajectories (D.3-12), whereas (D.3-11) for quick-

start technologies does not. The formulation presented here is based on the tight and compact formulation proposed in [70]. Furthermore, Morales-España et al. [71] has proven that the set of constraints (D.2-10)-(D.2-12) together with (D.3-9)-(D.3-12) is the tightest possible representation (i.e., convex hull) for the power-based model.

$$p_{\omega gt} + r^{+}_{\omega gt} \leq \left(\overline{P}_g - \underline{P}_g\right)u_{\omega gt} - \left(\overline{P}_g - SD_g\right)z_{\omega g,t+1} + \left(SU_g - \underline{P}_g\right)y_{\omega g,t+1} \quad \forall \omega, g, t \tag{D.3-9}$$

$$p_{\omega gt} - r^{-}_{\omega gt} \geq 0 \quad \forall \omega, g, t \tag{D.3-10}$$

$$\hat{p}_{\omega gt} = \underline{P}_g\left(u_{\omega gt} + y_{\omega g,t+1}\right) + p_{\omega gt} \quad \forall \omega, g \in \mathcal{G}^F, t \tag{D.3-11}$$

$$\hat{p}_{\omega gt} = \underbrace{\sum_{k=1}^{K_g}\sum_{i=1}^{SU^D_{gk}} P^{SU}_{gki}\delta_{\omega gk,\left(t-i+SU^D_{gk}+2\right)}}_{Startup\ trajectory} + \underbrace{\sum_{i=2}^{SD^D_g+1} P^{SD}_{gi} z_{\omega g,(t-i+2)}}_{Shutdown\ trajectory} + \underbrace{\underline{P}_g\left(u_{\omega gt} + y_{\omega g,t+1}\right) + p_{\omega gt}}_{Output\ when\ being\ up}$$

$$\forall \omega, g \in \mathcal{G}^S, t \tag{D.3-12}$$

ESS constraints for storage level (D.2-23) and storage level limits including reserve (D.2-24)-(D.2-25) continue the same. Nevertheless, the logic to avoid charging and discharging at the same time (D.3-13)-(D.3-14) is updated to consider the power output and charged power.

$$c_{\omega st} \leq (1 - \gamma_{\omega st}) \cdot \left(X^0_s + \overline{X}_s\right) \quad \forall \omega, s, t \tag{D.3-13}$$

$$\hat{p}_{\omega st} \leq \gamma_{\omega st} \cdot \left(X^0_s + \overline{X}_s\right) \quad \forall \omega, s, t \tag{D.3-14}$$

One of the main advantages of power-based formulation is that it allows to describe a more detailed set of constraints to represent the flexibility requirements, which are described in terms of power instead of energy. The proposed power-based equations in [70] ensure that reserves can be provided at any time within the hour by guaranteeing that the reserve does not exceed the ramp-capability at τ-min (e.g., τ=5 min) and power-capacity limits at the end of the hour (i.e., 60 min). Therefore, (D.3-15)-(D.3-16) guarantee that τ-min ramp capability is ensured for thermal technologies, while (D.3-17)-(D.3-18) guarantee the power-capacity limit for both τ-min and at the end of the hour. Although [70] shows the case for thermal technologies, we use the same concepts and extend the concept for ESS in (D.3-19)-(D.3-22).

$$\frac{\tau\left(p_{\omega gt} - p_{\omega g,t-1}\right)}{60} + r^{+}_{\omega gt} \leq \tau RU_g u_{\omega gt} \quad \forall \omega, g, t \tag{D.3-15}$$

$$\frac{\tau\left(p_{\omega gt} - p_{\omega g,t-1}\right)}{60} - r^{-}_{\omega gt} \geq -\tau RD_g u_{\omega g,t-1} \quad \forall \omega, g, t \tag{D.3-16}$$

$$\frac{\tau p_{\omega gt} + (60-\tau)p_{\omega g,t-1}}{60} + r^{+}_{\omega gt} \leq \left(\overline{P}_g - \underline{P}_g\right)u_{\omega gt} \quad \forall \omega, g, t \tag{D.3-17}$$

$$\frac{\tau p_{\omega gt} + (60-\tau)p_{\omega g,t-1}}{60} - r^{-}_{\omega gt} \geq 0 \quad \forall \omega, g, t \tag{D.3-18}$$

$$\frac{\tau\left(\hat{p}_{\omega st}-\hat{p}_{\omega s,t-1}\right)}{60}-\frac{\tau\left(c_{\omega st}-c_{\omega s,t-1}\right)}{60}+r^{+}_{\omega st} \leq \tau RU_g\left(X^0_s+x_s\right) \quad \forall \omega,s,t \tag{D.3-19}$$

$$\frac{\tau\left(\hat{p}_{\omega st}-c_{\omega st}\right)+(60-\tau)\left(\hat{p}_{\omega s,t-1}-c_{\omega s,t-1}\right)}{60}+r^{+}_{\omega st} \leq X^0_s+x_s \quad \forall \omega,s,t \tag{D.3-20}$$

$$\frac{\tau\left(\hat{p}_{\omega st}-\hat{p}_{\omega s,t-1}\right)}{60}-\frac{\tau\left(c_{\omega st}-c_{\omega s,t-1}\right)}{60}-r^{-}_{\omega st} \geq -\tau RD_g\left(X^0_s+x_s\right) \quad \forall \omega,s,t \tag{D.3-21}$$

$$\frac{\tau\left(\hat{p}_{\omega st}-c_{\omega st}\right)+(60-\tau)\left(\hat{p}_{\omega s,t-1}-c_{\omega s,t-1}\right)}{60}-r^{-}_{\omega st} \geq 0 \quad \forall \omega,s,t \tag{D.3-22}$$

REFERENCES

[1] O. J. Guerra, D. A. Tejada, and G. V. Reklaitis, "An optimization framework for the integrated planning of generation and transmission expansion in interconnected power systems," *Appl. Energy*, vol. 170, pp. 1–21, May 2016.

[2] D. A. Tejada-Arango, P. Sanchez-Martin, and A. Ramos, "Security Constrained Unit Commitment Using Line Outage Distribution Factors," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 329–337, Jan. 2018.

[3] O. J. Guerra, D. A. Tejada, and G. V. Reklaitis, "Climate change impacts and adaptation strategies for a hydro-dominated power system via stochastic optimization," *Appl. Energy*, vol. 233–234, pp. 584–598, Jan. 2019.

[4] International Energy Agency (IEA), "World Energy Outlook 2017," IEA, Paris, France, Technical Report, Nov. 2017.

[5] "White Paper - Grid integration of large-capacity Renewable Energy sources and use of large-capacity Electrical Energy Storage," International Electrotechnical Commission (IEC), 2012.

[6] J. Widén *et al.*, "Variability assessment and forecasting of renewables: A review for solar, wind, wave and tidal resources," *Renew. Sustain. Energy Rev.*, vol. 44, pp. 356–375, Apr. 2015.

[7] "CPC - Climate Weather Linkage: El Niño Southern Oscillation." [Online]. Available: http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/enso.shtml#history. [Accessed: 10-Aug-2015].

[8] "White Paper - Electrical Energy Storage," International Electrotechnical Commission (IEC), 2011.

[9] F. J. de Sisternes, J. D. Jenkins, and A. Botterud, "The value of energy storage in decarbonizing the electricity sector," *Appl. Energy*, vol. 175, pp. 368–379, Agosto 2016.

[10] D. A. Tejada-Arango, A. S. Siddiqui, S. Wogrin, and E. Centeno, "A Review of Energy Storage System Legislation in the US and the European Union," *Curr. Sustain. Energy Rep.*, vol. 6, no. 1, pp. 22–28, Jan. 2019.

[11] International Renewable Energy Agency (IRENA), "Planning for the renewable future: Long-term modelling and tools to expand variable renewable power in emerging economies." Jan-2017.

[12] J. Haas *et al.*, "Challenges and trends of energy storage expansion planning for flexibility provision in low-carbon power systems – a review," *Renew. Sustain. Energy Rev.*, vol. 80, pp. 603–619, Dec. 2017.

[13] J. Zhu, *Optimization of Power System Operation*. John Wiley & Sons, 2015.

[14] S. A.-H. Soliman and A.-A. H. Mantawy, *Modern Optimization Techniques with Applications in Electric Power Systems*. Springer Science & Business Media, 2011.

[15] M. Bertocchi, G. Consigli, and M. A. H. Dempster, Eds., *Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strategies*. New York: Springer-Verlag, 2011.

[16] A. A.-A. Ramos Antonio; Pérez, Gloria, *Optimización bajo incertidumbre*. Universidad Pontificia Comillas, 2008.

[17] PSR – Energy Consulting and Analytics, "SDDP – Stochastic hydrothermal dispatch with network restrictions," *Software | PSR – Energy Consulting and Analytics*, 2018. [Online]. Available: https://www.psr-inc.com/softwares-en/. [Accessed: 09-May-2018].

[18] Energy Exemplar, "PLEXOS® Integrated Energy Model," *Energy Exemplar » Energy Market Modelling*, 2018. [Online]. Available: https://energyexemplar.com/software/plexos-desktop-edition/. [Accessed: 09-May-2018].

[19] Birger Mo, "ProdRisk," *SINTEF*, 2018. [Online]. Available: http://www.sintef.no/en/software/prodrisk/. [Accessed: 09-May-2018].

[20] Andrés Ramos, "StarNet Model," *StarNet Model (Bulk Production Cost Model)*, 2018. [Online]. Available: https://www.iit.comillas.edu/aramos/starnet.htm. [Accessed: 27-Apr-2018].

[21] R. Moreno, R. Ferreira, L. Barroso, H. Rudnick, and E. Pereira, "Facilitating the Integration of Renewables in Latin America: The Role of Hydropower Generation and Other Energy Storage Technologies," *IEEE Power Energy Mag.*, vol. 15, no. 5, pp. 68–80, Sep. 2017.

[22] S. Rebennack, P. M. Pardalos, M. V. F. Pereira, and N. A. Iliadis, Eds., *Handbook of Power Systems I*. Berlin Heidelberg: Springer-Verlag, 2010.

[23] R. Billinton, R. Ringlee, and A. J. Wood, "Application to Generation Planning," in *Power-System Reliability Calculations*, MIT Press, 2003, pp. 46–100.

[24] K. Y. Lee and M. A. El-Sharkawi, "Applications to System Planning," in *Modern Heuristic Optimization Techniques:Theory and Applications to Power Systems*, Wiley-IEEE Press, 2008, pp. 750-.

[25] R. Loulou and M. Labriet, "ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure," *Comput. Manag. Sci.*, vol. 5, no. 1–2, pp. 7–40, Feb. 2008.

[26] "NREL: Energy Analysis - Regional Energy Deployment System (ReEDS) Model," 2016. [Online]. Available: http://www.nrel.gov/analysis/reeds/. [Accessed: 12-Apr-2016].

[27] "Resource Planning Model | Energy Analysis | NREL," 2016. [Online]. Available: https://www.nrel.gov/analysis/models-rpm.html. [Accessed: 30-Jan-2018].

[28] W. Lise, J. Sijm, and B. F. Hobbs, "The Impact of the EU ETS on Prices, Profits and Emissions in the Power Sector: Simulation Results with the COMPETES EU20 Model," *Environ. Resour. Econ.*, vol. 47, no. 1, pp. 23–44, Apr. 2010.

[29] R. Kannan, "The development and application of a temporal MARKAL energy system model using flexible time slicing," *Appl. Energy*, vol. 88, no. 6, pp. 2261–2272, Jun. 2011.

[30] R. Kannan and H. Turton, "A Long-Term Electricity Dispatch Model with the TIMES Framework," *Environ. Model. Assess.*, vol. 18, no. 3, pp. 325–343, Jun. 2013.

[31] G. Morales-España, L. Ramírez-Elizondo, and B. F. Hobbs, "Hidden power system inflexibilities imposed by traditional unit commitment formulations," *Appl. Energy*, vol. 191, pp. 223–238, Apr. 2017.

[32] G. Morales-España, R. Baldick, J. García-González, and A. Ramos, "Power-Capacity and Ramp-Capability Reserves for Wind Integration in Power-Based UC," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 614–624, Apr. 2016.

[33] "Renewables 2016 Global Status Report."

[34] L. Xie *et al.*, "Wind Integration in Power Systems: Operational Challenges and Possible Solutions," *Proc. IEEE*, vol. 99, no. 1, pp. 214–232, Jan. 2011.

[35] E. Lannoye, D. Flynn, and M. O'Malley, "The role of power system flexibility in generation planning," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–6.

[36] H. Zhao, Q. Wu, S. Hu, H. Xu, and C. N. Rasmussen, "Review of energy storage system for wind power integration support," *Appl. Energy*, vol. 137, pp. 545–553, Jan. 2015.

[37] P. Denholm and M. Hand, "Grid flexibility and storage required to achieve very high penetration of variable renewable electricity," *Energy Policy*, vol. 39, no. 3, pp. 1817–1830, Mar. 2011.

[38] Australian Energy Market Operator (AEMO), "Initial operation of the Hornsdale Power Reserve Battery Energy Storage System," Technical Report, Apr. 2018.

[39] H. Ibrahim, A. Ilinca, and J. Perron, "Energy storage systems—Characteristics and comparisons," *Renew. Sustain. Energy Rev.*, vol. 12, no. 5, pp. 1221–1250, Jun. 2008.

[40] "Issue Brief: A Survey of State Policies to Support Utility-Scale and Distributed-Energy Storage." National Renewable Energy Laboratory (NREL), 2016.

[41] G. Strbac *et al.*, "Opportunities for Energy Storage: Assessing Whole-System Economic Benefits of Energy Storage in Future Electricity Systems," *IEEE Power Energy Mag.*, vol. 15, no. 5, pp. 32–41, Sep. 2017.

[42] IRENA, "Electricity Storage and Renewables: Cost and Markets to 2030," International Renewable Energy Agency, Abu Dhabi, Oct. 2017.

[43] Sandia, "DOE/EPRI Electricity Storage Handbook in Collaboration with NRECA," Sandia National Laboratories, Albuquerque, USA, Technical Report SAND2015-1002, Feb. 2015.

[44] B. Zakeri and S. Syri, "Electrical energy storage systems: A comparative life cycle cost analysis," *Renew. Sustain. Energy Rev.*, vol. 42, pp. 569–596, Feb. 2015.

[45] V. Krishnan *et al.*, "Co-optimization of electricity transmission and generation resources for planning and policy analysis: review of concepts and modeling approaches," *Energy Syst.*, pp. 1–36, Aug. 2015.

[46] M. Boiteux, "Peak load-pricing," *The Journal of Business*, vol. 33, no. 2, pp. 157–179, 1960.

[47] E. Centeno, J. Reneses, and J. Barquin, "Strategic Analysis of Electricity Markets Under Uncertainty: A Conjectured-Price-Response Approach," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 423–432, Feb. 2007.

[48] S. Wogrin, P. Duenas, A. Delgadillo, and J. Reneses, "A New Approach to Model Load Levels in Electric Power Systems With High Renewable Penetration," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2210–2218, Sep. 2014.

[49] K. Eurek *et al.*, "Regional Energy Deployment System (ReEDS) Model Documentation: Version 2016," NREL (National Renewable Energy Laboratory (NREL), Golden, CO (United States)), 2016.

[50] S. Wogrin and D. F. Gayme, "Optimizing Storage Siting, Sizing, and Technology Portfolios in Transmission-Constrained Networks," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3304–3313, Nov. 2015.

[51] S. Wogrin, D. Galbally, and J. Reneses, "Optimizing Storage Operations in Medium- and Long-Term Power System Models," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–10, 2015.

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

[53] K. Poncelet, H. Höschle, E. Delarue, A. Virag, and W. D'haeseleer, "Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1936–1948, May 2017.

[54] Elaine Hale, Brady Stoll, and Trieu Mai, "Capturing the Impact of Storage and Other Flexible Technologies on Electric System Planning," National Renewable Energy Laboratory (NREL), Technical Report NREL/TP-6A20-65726, May 2016.

[55] S. Fazlollahi, S. L. Bungener, P. Mandel, G. Becker, and F. Maréchal, "Multi-objectives, multi-period optimization of district energy systems: I. Selection of typical operating periods," *Comput. Chem. Eng.*, vol. 65, pp. 54–66, Jun. 2014.

[56] M. S. ElNozahy, M. M. A. Salama, and R. Seethapathy, "A probabilistic load modelling approach using clustering algorithms," in *2013 IEEE Power Energy Society General Meeting*, 2013, pp. 1–5.

[57] S. Pineda and J. M. Morales, "Chronological time-period clustering for optimal capacity expansion planning with storage," *IEEE Trans. Power Syst.*, pp. 1–1, 2018.

[58] F. J. D. Sisternes, M. D. Webster, O. J. D. Sisternes, and M. D. Webster, "Optimal Selection of Sample Weeks for Approximating the Net Load in Generation Planning Problems," 2013.

[59] B. F. Hobbs, M. H. Rothkopf, R. P. O'Neill, and H. Chao, *The Next Generation of Electric Power Unit Commitment Models*. Springer Science & Business Media, 2006.

[60] B. S. Palmintier and M. D. Webster, "Impact of Operational Flexibility on Electricity Generation Planning With Renewable and Carbon Targets," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 672–684, Apr. 2016.

[61] E. Centeno, S. Wogrin, and A. Nogales, "Impact of technical operational details on generation expansion in oligopolistic power markets," *IET Gener. Transm. Distrib.*, vol. 10, no. 9, pp. 2118–2126, Jun. 2016.

[62] K. Poncelet, E. Delarue, D. Six, J. Duerinck, and W. D'haeseleer, "Impact of the level of temporal and operational detail in energy-system planning models," *Appl. Energy*, vol. 162, pp. 631–643, Jan. 2016.

[63] N. E. Koltsaklis and A. S. Dagoumas, "State-of-the-art generation expansion planning: A review," *Appl. Energy*, vol. 230, pp. 563–589, Nov. 2018.

[64] B. Hua, R. Baldick, and J. Wang, "Representing Operational Flexibility in Generation Expansion Planning Through Convex Relaxation of Unit Commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2272–2281, Mar. 2018.

[65] G. Morales-España, "Unit Commitment: Computational Performance, System Representation and Wind Uncertainty Management," Doctoral thesis, Universidad Pontificia Comillas, KTH Royal Institute of Technology, and Delft University of Technology, 2014.

[66] B. F. Hobbs, "Optimization methods for electric utility resource planning," *Eur. J. Oper. Res.*, vol. 83, no. 1, pp. 1–20, May 1995.

[67] V. Oree, S. Z. Sayed Hassen, and P. J. Fleming, "Generation expansion planning optimisation with renewable energy integration: A review," *Renew. Sustain. Energy Rev.*, vol. 69, pp. 790–803, Mar. 2017.

[68] H. Saboori and R. Hemmati, "Considering Carbon Capture and Storage in Electricity Generation Expansion Planning," *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1371–1378, Oct. 2016.

[69] B. F. Hobbs and S. S. Oren, "Three Waves of U.S. Reforms: Following the Path of Wholesale Electricity Market Restructuring," *IEEE Power Energy Mag.*, vol. 17, no. 1, pp. 73–81, Jan. 2019.

[70] G. Morales-Espana, A. Ramos, and J. Garcia-Gonzalez, "An MIP Formulation for Joint Market-Clearing of Energy and Reserves Based on Ramp Scheduling," *Power Syst. IEEE Trans. On*, vol. 29, no. 1, pp. 476–488, Jan. 2014.

[71] G. Morales-España, C. Gentile, and A. Ramos, "Tight MIP formulations of the power-based unit commitment problem," *Spectr.*, vol. 37, no. 4, pp. 929–950, Oct. 2015.

[72] I. Usera, P. Rodilla, S. Burger, I. Herrero, and C. Batlle, "The Regulatory Debate About Energy Storage Systems: State of the Art and Open Issues," *IEEE Power Energy Mag.*, vol. 15, no. 5, pp. 42–50, Sep. 2017.

[73] Pablo Rodilla *et al.*, "Energy Storage Perspectives: Challenges, Opportunities, and Future Directions," MIT Energy Initiative and Universidad Pontificia Comillas, 2013 ENI EES, Jan. 2016.

[74] A. J. Conejo and R. Sioshansi, "Rethinking restructured electricity market design: Lessons learned and future needs," *Int. J. Electr. Power Energy Syst.*, vol. 98, pp. 520–530, Jun. 2018.

[75] U.S. Energy Information Administration, "U.S. Battery Storage Market Trends," EIA, Washington, DC, Technical Report, May 2018.

[76] David Hart and Alfred Sarkissian, "Deployment of Grid-Scale Batteries in the United States," U.S. Department of Energy, Jun. 2016.

[77] H. Chen, S. Baker, S. Benner, A. Berner, and J. Liu, "PJM Integrates Energy Storage: Their Technologies and Wholesale Products," *IEEE Power Energy Mag.*, vol. 15, no. 5, pp. 59–67, Sep. 2017.

[78] FERC, "Electric Storage Participation in Markets Operated by Regional Transmission Organizations and Independent System Operators (Order 841)." Federal Energy Regulatory Commission (FERC), 15-Feb-2018.

[79] FERC, "Utilization of Electric Storage Resources for Multiple Services When Receiving Cost-Based Rate Recovery," Federal Energy Regulatory Commission, Washington, DC, Policy Statement PL17-2–000, Jan. 2017.

[80] "Decision on Multiple-use Application Issues." Public Utilities Commission of the State of California (CPUC), 17-Jan-2018.

[81] WIP and CENER, "European Regulatory and Market Framework for Electricity Storage Infrastructure. Analysis and recommendations for improvements based on a stakeholder consultation." stoRE Project, Jun-2013.

[82] "Commission proposes new rules for consumer centred clean energy transition - Energy - European Commission," *Energy*. [Online]. Available: /energy/en/news/commission-proposes-new-rules-consumer-centred-clean-energy-transition. [Accessed: 12-Jun-2017].

[83] "Lights and shadows for battery energy storage in the Winter Package - 30 November 2016 | Association of European Automotive and Industrial Battery Manufacturers." [Online]. Available: https://eurobat.org/lights-and-shadows-battery-energy-storage-winter-package-30-november-2016. [Accessed: 12-Jun-2017].

[84] European Commission, "COMMISSION STAFF WORKING DOCUMENT: Energy storage – the role of electricity." 02-Jan-2017.

[85] "Energy - Horizon 2020 - European Commission," *Horizon 2020*. [Online]. Available: /programmes/horizon2020/en/area/energy. [Accessed: 13-Jun-2017].

[86] "Strategic Energy Technology Plan - Energy - European Commission," *Energy*. [Online]. Available: /energy/en/topics/technology-and-innovation/strategic-energy-technology-plan. [Accessed: 13-Jun-2017].

[87] Inés Usera, Pablo Rodilla, Scott Burger, Ignacio Herrero, and Carlos Batlle, "State of the art and open issues in the regulatory debate about energy storage systems," Working Paper, 2017.

[88] R. Sioshansi, "Using Storage-Capacity Rights to Overcome the Cost-Recovery Hurdle for Energy Storage," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2028–2040, May 2017.

[89] A. Zucker, A. Hinchliffe, A. Spisto, European Commission, Joint Research Centre, and Institute for Energy and Transport, *Assessing storage value in electricity markets: a literature review*. Luxembourg: Publications Office, 2013.

[90] Ministerio de Energía, *Establece un nuevo sistema de transmisión eléctrica y crea un organismo coordinador independiente del sistema eléctrico nacional*. 2016.

[91] IEEE PES, *T&D 2018 Super Session: Reality Check on Energy Storage*. Chicago, Illinois, 2018.

[92] "Enel starts operations in UK at the Group's first stand-alone battery energy storage system." [Online]. Available: https://www.enel.com/media/press/d/2018/06/enel-starts-operations-in-uk-at-the-groups-first-stand-alone-battery-energy-storage-system-. [Accessed: 04-Oct-2018].

[93] "The role of storage in grid management." [Online]. Available: https://www.terna.it/en-gb/media/newsedeventi/ilruolodellostoragenellagestionedellereti.aspx. [Accessed: 04-Oct-2018].

[94] D. A. Tejada-Arango, S. Wogrin, and E. Centeno, "Representation of Storage Operations in Network-Constrained Optimization Models for Medium- and Long-Term Operation," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 386–396, Jan. 2018.

[95] D. A. Tejada-Arango, M. Domeshek, S. Wogrin, and E. Centeno, "Enhanced Representative Days and System States Modeling for Energy Storage Investment Analysis," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6534–6544, Nov. 2018.

[96] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, 1 edition. Hoboken, N.J: Wiley-Interscience, 2005.

[97] Q. Ploussard, L. Olmos, and A. Ramos, "An Operational State Aggregation Technique for Transmission Expansion Planning Based on Line Benefits," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2744–2755, Jul. 2017.

[98] E. Shayesteh, B. F. Hobbs, and M. Amelin, "Scenario reduction, network aggregation, and DC linearisation: which simplifications matter most in operations and planning optimisation?," *Transm. Distrib. IET Gener.*, vol. 10, no. 11, pp. 2748–2755, 2016.

[99] S. Inage, "Prospects for large-scale energy storage in decarbonised power grids," *Int. Energy Agency IEA*, 2009.

[100] ENTSO-E, "Ten Year Network Development Plan 2016," 2016. [Online]. Available: http://tyndp.entsoe.eu/. [Accessed: 03-Aug-2017].

[101] I. Staffell and S. Pfenninger, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy*, vol. 114, pp. 1224–1239, Nov. 2016.

[102] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251–1265, Nov. 2016.

[103] Diego A. Tejada-Arango, Sonja Wogrin, Afzal Siddiqui, and Efraim Centeno, "Short-term storage signals in hydrothermal dispatch models using a linked representative periods approach," *IIT-18-091*, Jul. 2018.

[104] N. E. Koltsaklis and M. C. Georgiadis, "A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints," *Appl. Energy*, vol. 158, pp. 310–331, Nov. 2015.

[105] L. Kotzur, P. Markewitz, M. Robinius, and D. Stolten, "Time series aggregation for energy system design: Modeling seasonal storage," *Appl. Energy*, vol. 213, pp. 123–135, Mar. 2018.

[106] L. Reichenberg, A. S. Siddiqui, and S. Wogrin, "Policy implications of downscaling the time dimension in power system planning models to represent variability in renewable output," *Energy*, vol. 159, pp. 870–877, Sep. 2018.

[107] J. M. Latorre, S. Cerisola, and A. Ramos, "Clustering algorithms for scenario tree generation: Application to natural hydro inflows," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1339–1353, Sep. 2007.

[108] W. L. de Oliveira, C. Sagastizábal, D. D. J. Penna, M. E. P. Maceira, and J. M. Damázio, "Optimal scenario tree reduction for stochastic streamflows in power generation planning problems," *Optim. Methods Softw.*, vol. 25, no. 6, pp. 917–936, Dec. 2010.

[109] A. Shapiro, "Analysis of stochastic dual dynamic programming method," *Eur. J. Oper. Res.*, vol. 209, no. 1, pp. 63–72, Feb. 2011.

[110] M. E. P. Maceira, V. Duarte, D. Penna, L. Moraes, and A. Melo, "Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil-description of the newave program," *16th PSCC Glasg. Scotl.*, pp. 14–18, 2008.

[111] V. R. Sherkat, R. Campo, K. Moslehi, and E. O. Lo, "Stochastic Long-Term Hydrothermal Optimization for a Multireservoir System," *IEEE Trans. Power Appar. Syst.*, vol. PAS-104, no. 8, pp. 2040–2050, Aug. 1985.

[112] S. Cerisola, J. M. Latorre, and A. Ramos, "Stochastic dual dynamic programming applied to nonconvex hydrothermal models," *Eur. J. Oper. Res.*, vol. 218, no. 3, pp. 687–697, May 2012.

[113] J. Reneses, J. Barquín, J. García-González, and E. Centeno, "Water value in electricity markets," *Int. Trans. Electr. Energy Syst.*, vol. 26, no. 3, pp. 655–670, 2016.

[114] M. V. F. Pereira and L. M. V. G. Pinto, “A Decomposition Approach to the Economic Dispatch of Hydrothermal Systems,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-101, no. 10, pp. 3851–3860, Oct. 1982.

[115] G. F. Bregadioli, E. C. Baptista, L. Nepomuceno, A. R. Balbo, and E. M. Soler, “Medium-term coordination in a network-constrained multi-period auction model for day-ahead markets of hydrothermal systems,” *Int. J. Electr. Power Energy Syst.*, vol. 82, pp. 474–483, Nov. 2016.

[116] J. Reneses, E. Centeno, and J. Barquin, “Coordination between medium-term generation planning and short-term operation in electricity markets,” *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 43–52, Feb. 2006.

[117] M. Maceiral *et al.*, “Twenty Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil-Main Features and Improvements on the NEWAVE Model,” in *2018 Power Systems Computation Conference (PSCC)*, 2018, pp. 1–7.

[118] B. G. Gorenstin, N. M. Campodonico, J. P. da Costa, and M. V. F. Pereira, “Stochastic optimization of a hydro-thermal system including network constraints,” *IEEE Trans. Power Syst.*, vol. 7, no. 2, pp. 791–797, May 1992.

[119] M. Alvarez, S. K. Rönnberg, J. Bermúdez, J. Zhong, and M. H. J. Bollen, “A Generic Storage Model Based on a Future Cost Piecewise-Linear Approximation,” *IEEE Trans. Smart Grid*, pp. 1–1, 2017.

[120] M. Alvarez, S. K. Rönnberg, J. Bermúdez, J. Zhong, and M. H. J. Bollen, “Reservoir-type hydropower equivalent model based on a future cost piecewise approximation,” *Electr. Power Syst. Res.*, vol. 155, pp. 184–195, Feb. 2018.

[121] D. Yacar, D. Tejada-Arango, and S. Wogrin, “Storage Allocation and Investment Optimization for Transmission Constrained Networks Considering Losses and High Renewable Penetration,” *IET Renew. Power Gener.*, vol. 12, no. 16, pp. 1949–1956, Oct. 2018.

[122] D. Gayme and U. Topcu, “Optimal power flow with large-scale storage integration,” *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 709–717, May 2013.

[123] Y. M. Atwa and E. F. El-Saadany, “Optimal Allocation of ESS in Distribution Systems With a High Penetration of Wind Energy,” *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1815–1822, Nov. 2010.

[124] A. Castillo and D. F. Gayme, “Profit maximizing storage allocation in power grids,” in *52nd IEEE Conference on Decision and Control*, 2013, pp. 429–435.

[125] A. Gopalakrishnan, A. U. Raghunathan, D. Nikovski, and L. T. Biegler, “Global optimization of multi-period optimal power flow,” in *2013 American Control Conference*, 2013, pp. 1157–1164.

[126] Z. Hu and W. T. Jewell, “Optimal power flow analysis of energy storage for congestion relief, emissions reduction, and cost savings,” in *2011 IEEE/PES Power Systems Conference and Exposition*, 2011, pp. 1–8.

[127] K. Purchala, L. Meeus, D. V. Dommelen, and R. Belmans, “Usefulness of DC power flow for active power flow analysis,” in *IEEE Power Engineering Society General Meeting, 2005*, 2005, pp. 454-459 Vol. 1.

[128] E. Sjödin, D. F. Gayme, and U. Topcu, “Risk-mitigated optimal power flow for wind powered grids,” in *2012 American Control Conference (ACC)*, 2012, pp. 4431–4437.

[129] M. Ghofrani, A. Arabali, M. Etezadi-Amoli, and M. S. Fadali, “A Framework for Optimal Placement of Energy Storage Units Within a Power System With High Wind Penetration,” *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 434–442, Apr. 2013.

[130] S. Lumbreras and A. Ramos, “The new challenges to transmission expansion planning. Survey of recent practice and literature review,” *Electr. Power Syst. Res.*, vol. 134, pp. 19–29, May 2016.

[131] M. F. Romlie, M. Rashed, C. Klumpner, and G. M. Asher, "An analysis of efficiency improvement with the installation of energy storage in power systems," in *7th IET International Conference on Power Electronics, Machines and Drives (PEMD 2014)*, 2014, pp. 1–6.

[132] F. Zhang, Z. Hu, and Y. Song, "Mixed-integer linear model for transmission expansion planning with line losses and energy storage systems," *Transm. Distrib. IET Gener.*, vol. 7, no. 8, pp. 919–928, Aug. 2013.

[133] Pedro Sánchez-Martín and Andrés Ramos, "Modeling Transmission Ohmic Losses in a Stochastic Bulk Production Cost Model," Oct. 1997.

[134] D. Z. Fitiwi, L. Olmos, M. Rivier, F. de Cuadra, and I. J. Pérez-Arriaga, "Finding a representative network losses model for large-scale transmission expansion planning with renewable energy sources," *Energy*, vol. 101, no. Supplement C, pp. 343–358, Apr. 2016.

[135] "14 Bus Power Flow Test Case," *Power Systems Test Case Archive*. [Online]. Available: https://www.ee.washington.edu/research/pstca/pf14/pg_tca14bus.htm. [Accessed: 26-Jun-2016].

[136] "Wind Integration Data Sets | Grid Modernization | NREL." [Online]. Available: https://www.nrel.gov/grid/wind-integration-data.html. [Accessed: 11-Dec-2018].

[137] European Association for Storage of Energy (EASE) and European Energy Research Alliance (EERA), "European Energy Storage Technology Development Roadmap Towards 2030," EASE/EERA, Technical Report, Mar. 2013.

[138] W. J. Cole, C. Marcy, V. K. Krishnan, and R. Margolis, "Utility-scale lithium-ion storage cost projections for use in capacity expansion models," in *2016 North American Power Symposium (NAPS)*, 2016, pp. 1–6.

[139] D. A. Tejada-Arango, G. Morales-España, S. Wogrin, and E. Centeno, "Power-Based Generation Expansion Planning for Flexibility Requirements," *ArXiv190207779 Math*, Feb. 2019.

[140] J. Meus, K. Poncelet, and E. Delarue, "Applicability of a Clustered Unit Commitment Model in Power System Modeling," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2195–2204, Mar. 2018.

[141] G. Morales-Espana and D. A. Tejada-Arango, "Modelling the Hidden Flexibility of Clustered Unit Commitment," *IEEE Trans. Power Syst.*, pp. 1–1, 2019.

[142] I. Momber, G. Morales-España, A. Ramos, and T. Gómez, "PEV Storage in Multi-Bus Scheduling Problems," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 1079–1087, Mar. 2014.

[143] G. Morales-Espana, J. M. Latorre, and A. Ramos, "Tight and Compact MILP Formulation of Start-Up and Shut-Down Ramping in Unit Commitment," *Power Syst. IEEE Trans. On*, vol. 28, no. 2, pp. 1288–1296, May 2013.

[144] C. Gentile, G. Morales-España, and A. Ramos, "A tight MIP formulation of the unit commitment problem with start-up and shut-down constraints," *EURO J. Comput. Optim.*, pp. 1–25, Apr. 2016.

[145] G. Morales-Espana, J. M. Latorre, and A. Ramos, "Tight and Compact MILP Formulation for the Thermal Unit Commitment Problem," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4897–4908, Nov. 2013.

[146] D. A. Tejada-Arango, "Case Studies for Power-based Capacity Expansion Planning: datejada/PB-CEP," 21-Dec-2018. [Online]. Available: https://github.com/datejada/PB-CEP. [Accessed: 21-Dec-2018].

[147] ENTSO-E, "Ten Year Network Development Plan 2018," 2018. [Online]. Available: https://tyndp.entsoe.eu/. [Accessed: 21-Dec-2018].

[148] H.-S. Park and C.-H. Jun, "A simple and fast algorithm for K-medoids clustering," *Expert Syst. Appl.*, vol. 36, no. 2, Part 2, pp. 3336–3341, Mar. 2009.

[149] M. Sun, F. Teng, X. Zhang, G. Strbac, and D. Pudjianto, "Data-Driven Representative Day Selection for Investment Decisions: A Cost-Oriented Approach," *IEEE Trans. Power Syst.*, pp. 1–1, 2019.

[150] A. Pyrgou, A. Kylili, and P. A. Fokaides, "The future of the Feed-in Tariff (FiT) scheme in Europe: The case of photovoltaics," *Energy Policy*, vol. 95, pp. 94–102, Aug. 2016.

[151] Iñigo Fernández de Araoz García-Miñaur, "Long-Term Investment In Energy Storage," Universidad Pontificia Comillas, Madrid, 2016.

[152] M. F. Anjos and A. J. Conejo, "Unit Commitment in Electric Energy Systems," *Found. Trends® Electr. Energy Syst.*, vol. 1, no. 4, pp. 220–310, Dec. 2017.

Article I

Article II

Article III

Article IV

Article V

Article VI

Article VII