COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES (GITI)

TRABAJO FIN DE GRADO (GITI)

SENSITIVITY ANALYSIS OF IMPACT OF CLUSTERIZATION METHODS IN MEDIUM-TERM POWER SYSTEM MODELS

Autor: Marta Niño Serrano
Directores: Sonja Wogrin
Diego A. Tejada Arango

Madrid
Junio de 2019

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESINAS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

La autora D. MARTA NIÑO SERRANO ser el titular de los derechos de propiedad intelectual de la obra: SENSITIVITY ANALYSIS OF IMPACT OF CLUSTERIZATION METHODS IN MEDIUM-TERM POWER SYSTEM MODELS, que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar "marcas de agua" o cualquier otro sistema de seguridad o de protección.
b) Reproducirla en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
e) Asignar por defecto a estos trabajos una licencia Creative Commons.
f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente)*.

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

a) Que la Universidad identifique claramente su nombre como autor de la misma

b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
c) Solicitar la retirada de la obra del repositorio por causa justificada.
d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que pudieran ejercitarse contra la Universidad por terceros que vieran infringidos sus derechos e intereses a causa de la cesión.
d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

6º. Fines y funcionamiento del Repositorio Institucional.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- La Universidad informará a los usuarios del archivo sobre los usos permitidos, y no garantiza ni asume responsabilidad alguna por otras formas en que los usuarios hagan un uso posterior de las obras no conforme con la legislación vigente. El uso posterior, más allá de la copia privada, requerirá que se cite la fuente y se reconozca la autoría, que no se obtenga beneficio comercial, y que no se realicen obras derivadas.
- La Universidad no revisará el contenido de las obras, que en todo caso permanecerá bajo la responsabilidad exclusive del autor y no estará obligada a ejercitar acciones legales en nombre del autor en el supuesto de infracciones a derechos de propiedad intelectual derivados del depósito y archivo de las obras. El autor renuncia a cualquier reclamación frente a la Universidad por las formas no ajustadas a la legislación vigente en que los usuarios hagan uso de las obras.
- La Universidad adoptará las medidas necesarias para la preservación de la

obra en un futuro.

- La Universidad se reserva la facultad de retirar la obra, previa notificación al autor, en supuestos suficientemente justificados, o en caso de reclamaciones de terceros.

Madrid, a ....5.... de ......Junio.................. de ...2019

ACEPTA

Fdo...

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título SENSITIVITY ANALYSIS OF IMPACT OF CLUSTERIZATION METHODS IN MEDIUM-TERM POWER SYSTEM MODELS en la ETS de Ingeniería – ICAI de la Universidad Pontificia Comillas en el curso académico 2018/2019 es de mi autoría, original e inédito y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

Fdo.: Marta Niño Serrano   Fecha: 5 / 6 / 2019

Autorizada la entrega del proyecto cuya información no es de carácter confidencial

LOS DIRECTORES DEL PROYECTO

Sonja Wogrin
Diego A.Tejada Arango

Fdo.: ……………… Fecha: 5 / 6 / 2019

Fdo.: ……………… Fecha: 5 / 6 / 2019

GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES (GITI)

TRABAJO FIN DE GRADO (GITI)

SENSITIVITY ANALYSIS OF IMPACT OF CLUSTERIZATION METHODS IN MEDIUM-TERM POWER SYSTEM MODELS

Autor: Marta Niño Serrano
Directores: Sonja Wogrin
Diego A. Tejada Arango

Madrid
Junio de 2019

SENSITIVITY ANALYSIS OF IMPACT OF CLUSTERIZATION METHODS IN MEDIUM-TERM POWER SYSTEM MODELS

Autor: Niño Serrano, Marta.

Directores: Wogrin, Sonja.

Tejada Arango, Diego A.

Entidad Colaboradora: IIT (Instituto de Investigación Tecnológica) – Universidad Pontificia Comillas

PROJECT ABSTRACT

In the past decades, electric power systems have been suffering severe changes due to the development of new technologies, such as renewable energy sources or storage facilities in form of batteries and hydro reservoirs. These energy sources and storage systems are key in the technological transformation we face nowadays. This kind of changes impact on the operation of power systems, whose behavior needs to be studied under new operating regimes. That is the reason behind the implementation of mathematical models that help us analyze all these issues and changes in power systems.

A standard and real problem faced by optimization models is the unit commitment (UC), which aims to schedule the most profitable generation type and units' combinations to meet the predicted demand. By this, each time slice in which the electric power system is reduced (traditionally hours), is assigned a generation level according to the imposed constraints. Due to the usual complexity of these systems, UC models are solved for a specific time horizon, typically a representative operation week.

Therefore, a whole year is 'reduced' to the representative operation week chosen prior to running the optimization problem itself by means of clusterization techniques and algorithms. These algorithms (kmeans and kmedoids) typically group several hours with a number of normalized characteristics (solar energy, demand, wind energy, etc). In this work, this type of clusterization is compared to a method including hierarchical clustering (Chronological model) which enables the possibility of including additional conditions on how clusters are merged or divided. Additionally to the clustering techniques used to reduce the employed time in running the operation problem, a shortened version of the unit commitment is used in order to evaluate magnitude of the generation mix scheduled. This simplification is based on the use of a relaxed modelling approach (rMIP) when obtaining the most cost-effective generating units for every hour of the time horizon.

In order to evaluate the impact of obtaining a relaxed solution or/and the use of one or another clusterization technique, several analyses are carried out in a yearly base

case, with a changeable RES production and with a diverse behavior from the BESS employed.

The solutions obtained through a rMIP programming typically make a 1-2% error in total energy production but entail a CPU time reduction of around 100 to 5000 (times faster). The combination of this approach with the clusterization techniques previously mentioned would mean, in absolute terms, a much smaller time commitment (nearly immediate solutions) with an intermediate error around 3%.

Moreover, the changeable renewable energy sources production is relevant in order to evaluate how these models behave with diverse generation mixes and because the RES penetration in Spain aims to increase in the following years in both Wind and Solar generation. The results throw light on a less cost-efficient function if the renewable production only depends on one generation type (Solar or Wind) being the Solar energy more expensive since it involves a higher use of storage systems that may saturate (BESS technologies). This is resumed in Figure 1.

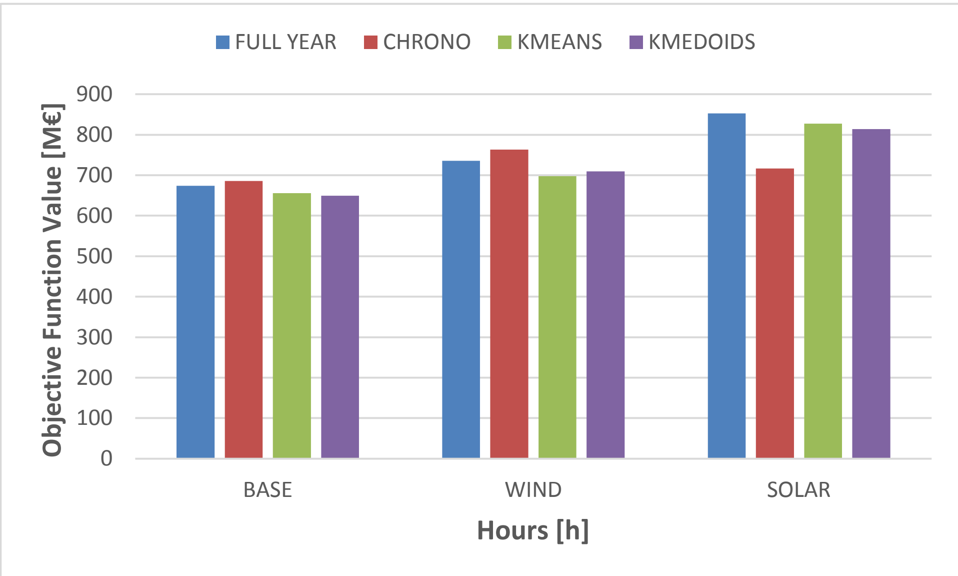


Figure 1. Objective Function for the three RES cases.

Lastly, a study of BESS performance is conveyed because this technology is likely to experiment an increase in its use as more renewable energies are introduced to the power system. BESS behavior is therefore studied through a changeable duration of its work cycle by modifying the ramps (up and down) that can take place. If work cycles are optimized (BESS do not saturate and spend the maximum number of hours charging and discharging energy), the total system costs will be reduced to its minimum, which happens for a duration of 8 hours (4 charge and the same discharging), as illustrated in Figure 2.

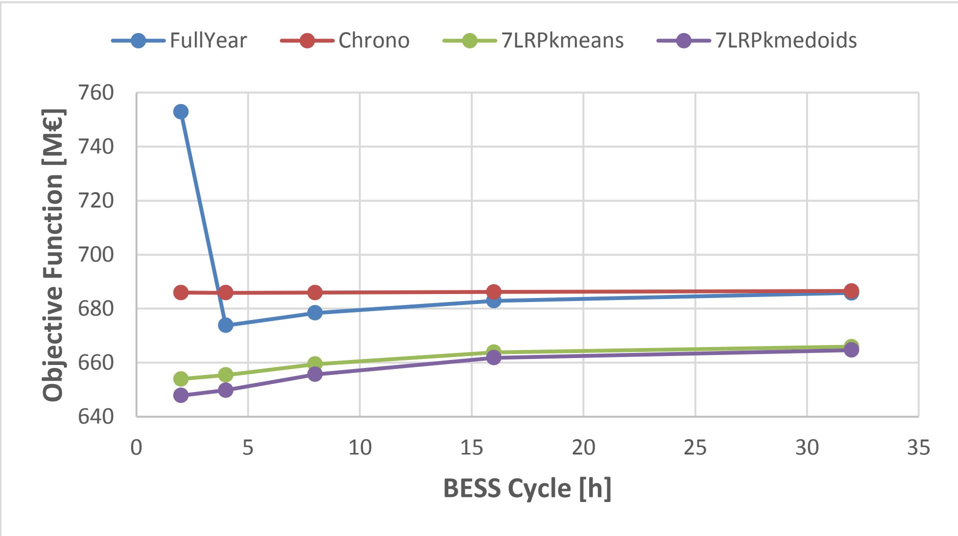


Figure 2. Objective function value – BESS charge cycle duration

ANÁLISIS DE SENSIBILIDAD DEL IMPACTO DE MÉTODOS DE CLUSTERIZACIÓN EN OPERACIÓN SISTEMAS ELÉCTRICOS

En las últimas décadas, los sistemas de energía eléctrica han sufrido grandes cambios debidos al desarrollo de nuevas tecnologías con la incorporación de fuentes de energía renovable y sistemas de almacenamiento en forma de baterías y reservas hidráulicas. Tanto estas nuevas fuentes de energía como los sistemas de almacenamiento de esta son claves en la transformación tecnológicas que se está produciendo actualmente. Este tipo de cambios tienen un impacto directo en la operación de los sistemas de energía eléctrica, cuyo comportamiento necesita ser estudiado bajo estos nuevos métodos de operación. Esta es la razón detrás del uso de modelos matemáticos que ayudan a analizar estas cuestiones y cambios en los sistemas eléctricos.

Un problema estándar (y real) al que se enfrentan los modelos de optimización es el Unit Commitment (UC), cuyo objetivo es la planificación del escenario de generación más rentable para cubrir la demanda pronosticada. Con esto, cada rebanada temporal en la que están divididos los sistemas (tradicionalmente horas), es asignada con un nivel de generación que cumple las restricciones impuestas. Dada la complejidad habitual de estos sistemas, modelos UC son resueltos para un horizonte temporal determinado, típicamente una semana representativa.

Por lo tanto, un año completo será reducido a una semana representativa de operación, seleccionada antes de proceder a correr el problema de optimización, mediante técnicas y algoritmos de clusterización. Estos algoritmos (kmeans y kmedoids) típicamente agrupan horas con un número determinado de características normalizadas (energía solar, demanda, energía eólica, etc). En este trabajo, este tipo de clusterización es comparada con un método que utiliza una clusterización jerárquica (modelo cronológico) que posibilita la inclusión de condiciones adicionales de cómo agrupar o dividir cada grupo o clúster. Adicionalmente al uso de técnicas de clusterización para reducir el tiempo empleado en resolver el problema de operación, una versión reducida del Unit Commitment es utilizada para evaluar el orden de magnitud del mix energético programado. Esta simplificación está basada en el uso de un modelado relajado (rMIP) a la hora de obtener las unidades de generación más rentables para cada hora del horizonte temporal.

Para evaluar el impacto de obtener una solución relajada y/o el uso de una u otra técnica de clusterización, diferentes análisis se han llevado a cabo en el caso base anual, con una producción renovable cambiante y con un comportamiento diferente en las baterías utilizadas.

La solución obtenida con una programación rMIP comete habitualmente un error del orden de 1-2% en la energía total producida, pero conlleva una reducción del tiempo de computación de entre 100 y 5000 veces menor.

La combinación de esta aproximación con las técnicas de clusterización anteriormente mencionadas significará, en términos absolutos, un empleo temporal mucho menor (soluciones casi inmediatas) con un error compromiso alrededor del 3%.

Además, una producción de energías renovables cambiante es relevante para evaluar cómo se comportan estos modelos con diferentes escenarios de generación y porque la penetración de energías renovables en España tenderá a aumentar en los próximos años tanto en energía eólica y producción solar.

Los resultados arrojan luz sobre una función objetivo menos rentable si la producción renovable depende sólo de un tipo de generación (Solar o energía eólica), siendo la energía solar más cara dado que implica un uso de los sistemas de almacenamiento, que a su vez pueden saturar (baterías). Todo esto viene resumido en la Figura 1.

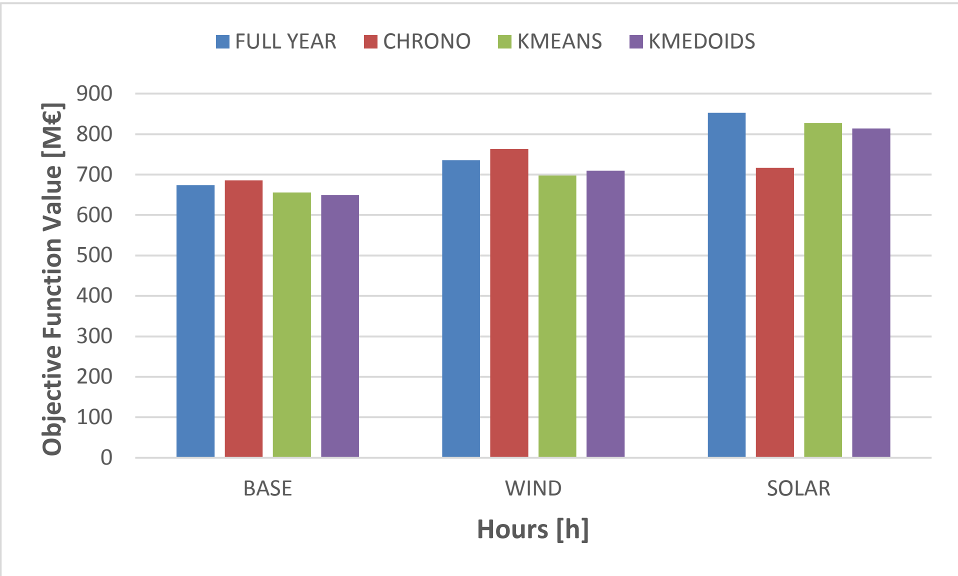


Figura 1. Función Objetivo para los tres casos renovables.

Por último, se ha llevado a cabo un estudio del funcionamiento de los sistemas BESS porque esta tecnología es probable que experimente un aumento en su utilización desde que un mayor porcentaje de renovables se introducirá en el sistema. El comportamiento de tecnologías BESS es, por tanto, estudiada mediante una duración de su ciclo de trabajo también cambiante. Esto se consigue modificando las rampas de subida y bajada que pueden darse. Si los ciclos de trabajo se optimizan (BESS no saturan y emplean el máximo número de horas cargando y descargando energía), el coste total del sistema se reducirá a su mínimo, lo cual sucede para una duración de 8 horas (y cargando y otras 4 descargando su energía), como se ilustra en la Figura

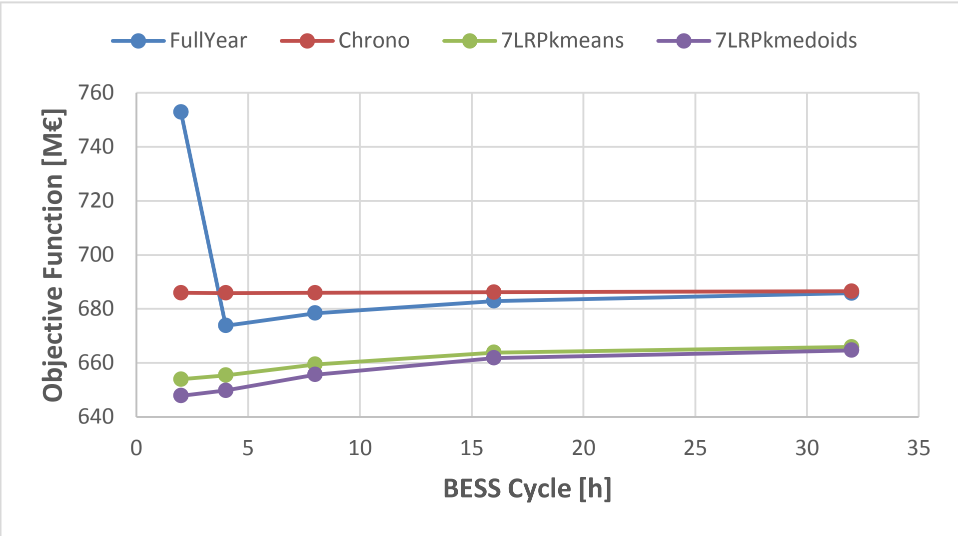


Figura 2. Función Objetivo- Duración del ciclo de carga de BESS

Index

1. Introduction 2
   a. Motivation 2
   b. Literature review 3
   c. Objectives 4
   d. Thesis Outline 5
2. Methodology 6
   a. Model Formulation 6
3. Base Case 12
   a. Base Case Results 12
4. Intra-Model Comparison 16
   a. Full Year 16
   b. Chrono 18
   c. 7LRP kmeans 19
   d. 7LRP kmedoids 21
5. Transversal Intra Model Comparison 22
6. Transversal Inter Model MIP Comparison 26
7. RES Sensitivity Analysis 30
   a. Wind 32
   b. Solar 35
8. Transversal BESS Sensitivity Analysis 38

References 46

Appendix 1 - Notation of Models 48

Indexes and sets 48

Parameters 48

Variables 48

Appendix 2 – Generation mix 50

Appendix 3 – rMIP Tables and Figures 54

rMIP General RES Sensitivity Analysis 54

rMIP WIND Sensitivity Analysis 55

rMIP SOLAR RES Sensitivity Analysis 57

rMIP BESS Sensitivity Analysis 59

Appendix 4 – Tools used 62

List of Figures

Figure 1. Objective Function for the three RES cases. IV
Figure 2. Objective function value – BESS charge cycle duration V
Figure 3. Equations (2w) and (2x): interactions between binary variables. 10
Figure 4. Generation mix for full year MIP model. 12
Figure 5. Detail of full year MIP model generation mix 13
Figure 6. Hydro evolution for full year MIP model 14
Figure 7. BESS evolution of a single week for MIP full year model 14
Figure 8. Hydro Evolution for Full Year Model 17
Figure 9. BESS Evolution Comparison of a single week for the Full Year Model 17
Figure 10. Hydro Evolution Comparison for the Chrono Model 18
Figure 11. BESS Evolution Comparison of a single week for the Chrono Model 19
Figure 12. BESS Evolution for 7LRP kmeans: MIP and rMIP. 20
Figure 13. Hydro Evolution Comparison for the 7LRP kmedoids Model 21
Figure 14. Price Error [%] in each model 23
Figure 15. Price evolution Inter-Model 26
Figure 16. Hydro levels Inter-Model. 27
Figure 17. Objective Function for the three RES cases. 30
Figure 18. RES Spillage [%] 31
Figure 19. Inter Model Comparison: BESS Level 33
Figure 20. BESS Evolution: 7LRP kmeans RES Wind production. 34
Figure 21. Hourly price evolution: RES Wind Production. 34
Figure 22. Hourly price evolution: RES Solar Production 35
Figure 23. BESS Evolution: 7LRP kmeans RES Solar production. 36
Figure 24. Transversal RES Analysis. 37
Figure 25.7LRP kmeans BESS level for work cycles of 8h and 16h. 38
Figure 26. 7LRP kmeans BESS level for work cycle of 32h. 39
Figure 27. Objective function value – BESS charge cycle duration 40
Figure 28. Hourly cost in BESS Analysis[€/MWh] 41
Figure 29. Hourly BESS Price Evolution for all Models in the selected period. 42
Figure 30. BESS level for a 16h-work cycle. 43
Figure 31. BESS level for a 32h-work cycle. 43
Figure 32. Objective Function for the three RES cases. 54
Figure 33. RES Spillage. 54
Figure 34. Inter Model Comparison: BESS Level 55
Figure 35. BESS Evolution: 7LRP kmeans RES Wind production. 56
Figure 36. Hourly price evolution: RES Wind Production. 56
Figure 37. Hourly price evolution: RES Solar Production 57
Figure 38. BESS Evolution: 7LRP kmeans RES Solar production. 58
Figure 39. rMIP Transversal RES Analysis. 58
Figure 40.7LRP kmeans BESS level for work cycles of 8h and 16h. 59
Figure 41. Objective function value – BESS charge cycle duration 59
Figure 42. Hourly cost in BESS Analysis[€/MWh] 60

List of Tables

Table 1. Summary of the time indices. 7
Table 2. Main output for full year MIP model 12
Table 3. Summary for Full Year model: MIP and rMIP 16
Table 4. Summary for Chrono model: MIP and rMIP 18
Table 5. Summary for 7LRP kmeans model: MIP and rMIP 19
Table 6. Summary for 7LRP kmedoids model: MIP and rMIP 21
Table 7. Comparison of Objective Function Error and CPU time ratio. 22
Table 8. Summary of errors in a rMIP approach by technology type. 23
Table 9. Summary of Model comparison 26
Table 10. Error in Inter-Model Generation Comparison 27
Table 11. Model and Case objective function Error 30
Table 12. (Best) Ordered models by RES generation mix 31
Table 13. Summary of results: RES Wind production. 32
Table 14. Inter Model Comparison: RES Wind Production. 33
Table 15. Summary of results: RES Solar production 35
Table 16. Inter-Model Comparison: RES Solar Production. 36
Table 17. Model and Case objective function Error 54
Table 18. Summary of results: RES Wind production. 55
Table 19.Inter Model Comparison: RES Wind Production. 55
Table 20. Summary of results: RES Solar production 57
Table 21. Inter-Model Comparison: RES Solar Production. 57

1. Introduction

a. Motivation

Over the past decades, electric power systems have evolved in a completely different way as opposed to how they traditionally did. The use of new technologies, especially renewable sources of energy and the correlated development of storage units are key in the technological transformation of the electric sector.

These changes in the generation sources of energy have logically had a great impact on the operation of the power system itself. Traditionally, thermal units conquered the generation scenario due to the high flexibility they hold, yet the introduction of emission targets for decarbonization in several countries over the world has reduced its presence in power systems. Additionally, the intermittency renewable generation implies, introduces uncertainty into the system, which stands against a rather predictable demand.

These are reasons to study the behaviour of electric power systems under different operation regimes through viable mathematical models that capture the planning horizon through time periods- typically hours.

Power system planning models include high resolution short-term models with hourly information that typically consider unit commitment to operate the system itself; and long-term models based on investment decisions completely unaware of small-time scale changes in order to reduce the computational burden that may imply the hourly time representation of the planning horizon.

Considering a complex energy system-with different generation and storage technologies- obliges to include both approaches: both a long-term system acting as benchmark and the short-term dynamics to account for hourly constraints.

As previously mentioned, the introduction of renewable energy sources (RES), with a high variability of the primary source of energy: i.e.: sunlight in photovoltaics or wind, makes it necessary to assess its treatment.

The potential clear solution to these types of renewable generation sources comes within energy storage systems (ESS). Dealing with these storage solutions, makes it necessary to have chronological information in consecutive hours or time periods and this is because the total stored energy at any hour depends on the charged or discharged energy of previous time periods.

The duration itself of charge and discharge cycles establishes a frontier between the type of storage used: while Renewable Energy Sources (RES) need to incorporate intra-day storage due to fast cycles, elements or technologies with higher inertia are considered in a lower time grid and need an inter-period storage.

This is the reason behind why ESS need to bond together with other types of storage units in the same scenario.

Matching immediate storage facilities such as Li-iOn batteries with longer period types of storage, such as Hydro power units, reveals one of the problems dealt in models including different time dynamic technologies: selecting the duration of the time slices and its weight.

Modelling all types of storage units in an hourly benchmark would render the model computationally intractable so in recent literature, clustering techniques are applied to reduce the amount of redundant information until the minimum representative data is reached.

In this thesis, three medium-term optimization models are proposed, based in [1], [2] and [3], so that they can model different time dynamics maintaining chronology to optimize the coexistence of all storage technologies.

Particularly, clustering techniques are used to accelerate computation time while maintaining details of the model itself. A compromise solution between those two ends of the scale-computational speed and detailed model information- is one of the key points of this thesis.

These algorithms (kmeans and kmedoids) typically group several hours with a number of normalized characteristics (solar energy, demand, wind energy, etc). In this paper this type of clusterization is compared to a method including hierarchical clustering (Chrono model) which enables the possibility of including additional conditions on how clusters are merged or divided. It also includes a measure of dissimilarity between two groups of observations (as well as a linkage criterion to determine which clusters are merged or divided in agglomerative or divisive hierarchical clustering).

b. Literature review

Short-term optimization techniques normally face the unit commitment (UC) problem. As seen in [4], unit commitment first works began in the 1950s, under the development of M. Boiteux and it schedules the most cost-effective generating unit combination to meet demand requirements. Each hour, units meet a generation level according to transmission network constraints and own generation limits.

While in their firsts developments, different generation sources were included, such as thermal of hydro pumping units, electric power systems started to get bigger and this hourly UC became computationally difficult to solve. Since nowadays UC models include very detailed information for every hour of the time horizon, they are naturally computationally complex to solve so they are limited to a specific time range.

On the other hand, long-term optimization models are investment-focused and err of being uncomplicated mathematical models since they do not include small-time scale changes.

In order to create feasible and relatively fast optimization models, the reduction of input data to the model is key. This is done through clustering algorithms, able to aggregate similar time periods (hours) in this case; although these techniques are commonly used in many other fields apart from power systems.

K-means algorithm is probably the simplest and most used clustering technique to aggregate typical periods, although k-medoids is commonly used in Mixed Integer Problems (MIP). Similarity criteria is also used to join 'adjacent' periods in groups in Hierarchical Clustering. All these methods are applied by preserving at the same time the chronology within periods.

Previous studies focus on either battery storage (short term) or hydro reservoirs (long term); though more recent studies as [1] and [2] have proposed models that deal with both ESS technology types.

c. Objectives

The work presented in this thesis is the extension of the developments presented in [1], [2] and [3]. The base objective is to extend the work of the mentioned approaches and broaden in the impact of changes in the input data to the model, through the application of aggregation techniques (formation of clusters).

Starting from the concepts described in those developments, this thesis will include a generalization in formulation of the model proposed, new ideas and several sensitivity analyses in the change in input data.

The objectives of this work are divided in three main points:

The first block is focused on evaluating the accuracy and computational efficiency of the different approaches to the model: whether a linear model or a MIP is considered. Some tests are developed in order to:

a. Determine the main solutions differences between the model approach and resolution. We also carry out an analysis of computational time employed, complexity of the approach and average total calculated error
b. Determine under which scenarios one approach is recommended over the alternative one.

The second key point and objective correspond to the analysis of the change in the input data in the model:

a. Asses the reliability and impact of using an hourly dispatch model, a chronological clustered model or the newly developed linked representative periods (LRP) model.

The third objective corresponds to sensitivity analyses carried out to test the developed methodology.

a. Analyze the impact of the change in renewable penetration to the model, considering either wind or solar generation in the last case.
b. Evaluate the impact on changing the Battery Energy Storage Systems (BESS) work cycle duration by maintaining its capacity fixed.

Through case studies a numerical analysis will be carried out in order to evaluate the performance of the model approaches and input data.

These three objectives will help in the growth of a robust system under different generation scenarios and the storage technologies included, able to handle predictions and forecasts for the upcoming years.

d. Thesis Outline

This document is organized as follows: Section 2.a includes a detailed description of the methodology employed in this project, which includes all formulations of mathematical optimization models, and a description of the three models evaluated in this thesis. while further sections include the explanation of how the two models that use data clustering work, with its explanation.

Results begin with Section 3.a, in which the main models results' are presented so that in further Section 4 every model has a subsection corresponding to the FullYear, Chrono and 7LRP kmeans and kmedoids intra model comparison. This translates into the implementation of a programming approach comparison (studying whether a relaxed programming approach is sufficiently accurate and fast to solve the optimization problem effectively compared to the real MIP programming) in Subsections 4.a, 4.b, 4.c and 4.d, respectively. Section 5 compiles a transversal comparison that coincides with the overview of the previous section.

On the contrary, and to complement an intra-model approach comparison, a contrast between all models is done in Section 6, although not doing so extensively since previous work in [2] made that already.

Further Sections 7 and 8 are dedicated to the study of the two main sensitivity analyses carried out in this document. While Section 7 explores the models behavior's when having only wind (Section 7.a) renewable generation or exclusively depending on solar renewable generation (Section 7.b), Section 8 is dedicated to a profound exploration in the BESS technology cycle duration.

2. Methodology

This section includes all the conceptual explanations of the different features that compound the Full Year, Chrono and 7LRP models. Section 2.a includes explanations towards the model's formulation, with its related indices, parameters, variables and equations. Although explanations are included in this part of the document, the following section includes a general scheme of the notation used in the optimization problem.

a. Model Formulation

This section includes the formulation of the hourly unit commitment model which in fact represents the exact solution to the optimization problem exposed. This model is also referred as Full Year MIP or base case, for further sections in this document. This model (FullYear) with a unit commitment approach (MIP programming) represents the system against which models -both Chrono and 7LRP- and approaches (RMIP or MIP) are compared. Since the models' formulation has the same structure in constraints' terms, all models' formulations are grouped and so when a specific parameter changes between the hourly unit commitment model, chronological one and the one embodying representative days of the yearly scenario, it can be explained.

For every hour, this model meets the optimal and precise generation solutions considering all the constraints listed below in this section. The main drawback of this model and approach is the computational burden it implies due to the large amount of information, equations and variables that need to be computed in the energy dispatch. It is precisely this computational tractability the reason why large-scale, detailed models with unit commitment cannot be solved over a long-term time horizon.

Hence approximations in the programming approach are considered, taking an rMIP approach as a faster way in obtaining the optimal solution. In the same way, a reduction of calculations is also done in more sophisticated models -Chrono and 7LRP in this document.

Firstly, the indices involved are explained. The index t represents all hours along the time horizon, which in this case is an entire year. For all models this index takes a value of 8736 hours of an entire year. Here it is necessary to describe two additional time indices, which are rp and k.

Index rp represents the number of representative periods considered in the model. For FullYear this index is 1 since it is a whole year the period considered representative. Since one of the objectives of this document is the comparison of both computationally reduced Chrono and 7LRP models, these indices are strategically chosen to do a fair comparison with both of them considering a representative week of the whole year as base for the remaining hours to be placed in either one cluster or another. Therefore index rp is 1 (1 representative week) for the Chrono model whether as for the 7LRP this index takes a value of 7 (7 representative days).

The 7LRP models considers rp to have a value of 7 representative days in the year with whom each hour of the studied year will be placed into one day or another according to the 'distance' the clusterization algorithm calculates -being these algorithms kmeans and kmedoids

The following index k is directly related to the previous one since it represents the number of periods inside a representative period. For the Chrono model, since one representative week wants to be considered, 168 hours are the ones corresponding to this formulation while the number of hours in 7LRP are 24 as there have been defined 7 representative days in the week.

Since this formulation may have not been sufficiently clear, Table 1 sums this information.

MODEL	FULL YEAR	CHRONO	7LRP
rp	1	1	7
k [h]	8736	168	24
t [h]	8736	8736	8736
Representative hours [h]	8736 (1 year)	168 (1 week)	168 (1 week)

Table 1. Summary of the time indices.

The generation technology type is described with the index j, although subindices $g \subseteq \mathcal{J}$, $s \subseteq \mathcal{J}$ and $v \subseteq \mathcal{J}$ are used to precise the type of technology of the generating unit, being thermal generators, storage units or renewable sources, respectively. Subindex $g \subseteq \mathcal{J}$, characterizes the thermal generating units such as coal, combined cycle or nuclear energy. On the other hand, storage such as hydro units or battery energy storage systems (BESS) are under the subindex $s \subseteq \mathcal{J}$. Lastly, subindex $v \subseteq \mathcal{J}$ is used to represent wind and solar energy, both included in a renewable generation scenario.

Index i represents a node of the power system, which in our case is only Spain. If a generalization of the system is required, node i could represent connected countries around Europe, as an example. This index is representative for the definition of g_i, being this g_i the corresponding generators connected to node i, Spain. As nodes are defined, lines connecting pairs are presented with the index $l \in \mathcal{L}$, having both output and input lines to a specific node i; $lo \subseteq \mathcal{L}$ are output lines whereas $li \subseteq \mathcal{L}$ are input lines.

Let us now present the needed models' parameters. In relation to the pretended costs' reduction, all costs need to be determined. Therefore, each generation technology j will have three related costs which correspond to a fix cost of the unit startup: C_j^{st} and two variable costs of the unit being connected or not to the production system: C_j^{int} and to the power produced in each unit: C_j^{var}. In the variable terms, C_j^{int} has units of cost per hour connected to the system whether C_j^{var} dictates the total cost per GWh produced.

In order to do a fair comparison between clustered models, a representative periods' and hours' weight is gathered with parameters W_{rp}^{rp} and W_{rp}^{k} that naturally represent the weight of each cluster in hours. In addition, another hour parameter is included, and it is M, a so-called moving window -with hour units- so that storage can be restrained.

Moreover, the system operation is based on the hourly match of power demand and production so a parameter gathering the hourly demand of the power system is included: $D_{rp,k,i}$. This parameter is defined by both time indices and the single node i of the Spanish system. In relation to the production that meets the hourly demand, a peak and valley generation of each technology needs to be defined again with parameters P_j^+ and P_j^- . Of course, there is essential to stablish a second level production limit corresponding to the maximum ramp up and down of each unit: RU_j and RD_j- that could produce from one hour to the following one. It is important to determine that these production parameters do not have a perfect efficiency, so are logically affected by a E_s pumping reduction efficiency parameter, as well as the possible interruption of the thermal generation units (as not all units produce every hour of the year; i.e.: maintenance shutdowns…) are modelled with ε_g parameter.

Lastly, there exist inflows for the hydro storage, annexed with parameter $D_{rp,k,s}$, only specified to work for hydro storage.

Lastly, let us present the variables and their dependencies with subindices of our model. Since our model includes a unit commitment modelling approach, some binary decisions are included with these types of variables. The connection of the generation unit in each representative period and hour is well modelled with the variable $u_{rp,k,g}$. The startups and shutdowns of the units are included with the binary variables $y_{rp,k,g}$ and $z_{rp,k,g}$, respectively. Remember in this part that there exists a single cost regarding all these connections, and productions that will be logically multiplied by the binary variables just explained.

Regarding the hourly balance demand and production equation, the unit production is included by means of the variable $p_{rp,k,j}$. The balance must be met in each node of the system so, although in this case there are no other nodes than Spain, a variable considering the potential power flows to the node is also included in case of wanting to broaden the formulation to several countries- $pf_{rp,k,l}$. Specifically targeting the storage evolution and consumption, there are variables of hourly storage consumption $q_{rp,k,s}$, spillage $sp_{rp,k,s}$ and reserve at two different inter and intra periods: $R_{rp,k,j}^{inter}$ and $R_{rp,k,j}^{intra}$ (there is also an initial reserve for the storage: $R0_s$ to force hydro levels to depart and finish at the same point for the first and last hour of the year). For the production of each thermal unit, the variable $\hat{p}_{rp,k,g}$ quantifies the production of each thermal generator above the minimum or its production floor). As it may occur, although not desired, the non-served power is modelled in the system formulation thanks to the variable $pns_{rp,k,i}$.

Hereunder, equations will be presented. They are outlined from general to ones. Constraints that involve all power units are first, thermal unit specifications are considered after and hydro plants in the end.

While equation (2a) embodies the system's objective function, equations between (2b) and (2c) describe all the constraints related to thermal power plants while equations (2d) and (2e) belong to plants that have energy storage capabilities (either BESS or hydro resources).

$$\begin{aligned} min \sum_{rp,k,i} W_{rp}^{rp} * W_k^k * C^{ENS} * pns_{rp,k,i} + \sum_{rp,k,s} W_{rp}^{rp} * W_k^k * \frac{C^{ENS}}{2} * sp_{rp,k,s} \\ + \sum_{rp,k,g} W_{rp}^{rp} * W_k^k * C_g^{st} * y_{rp,k,g} + \sum_{rp,k,g} W_{rp}^{rp} * W_k^k * C_g^{int} * u_{rp,k,g} \\ + \sum_{rp,k,g} W_{rp}^{rp} * W_k^k * C_g^{var} * p_{rp,k,g} \end{aligned} \quad \textbf{(2f)}$$

$$\sum_{g,i \in gi} p_{rp,k,j} - \sum_{s,i \in gi} \frac{q_{rp,k,s}}{E_s} + \sum_{l \in li} pf_{rp,k,l} - \sum_{l \in lo} pf_{rp,k,l} + pns_{rp,k,i} = \mathrm{D_{rp,k,i}} \quad \forall rp,k,a \quad (2g)$$

$$\hat{p}_{rp,k,g} \leq (P_g^+ - P_g^-) * (u_{rp,k,g} - y_{rp,k,g}) \quad \forall\, rp,k,g \quad (2h)$$

$$\hat{p}_{rp,k,g} \leq (P_g^+ - P_g^-) * (u_{rp,k,g} - z_{rp,k+1,g}) \quad \forall\, rp,k,g \quad (2i)$$

$$p_{rp,k,g} = u_{rp,k,g} * P_g^- + \hat{p}_{rp,k,g} \quad \forall\, rp,k,g \quad (2j)$$

$$u_{rp,k,g} - u_{rp,k-1,g} = y_{rp,k,g} - z_{rp,k,g} \quad \forall\, rp,k,g \quad (2k)$$

$$-u_{rp,k-1,g} * RD_g \leq \hat{p}_{rp,k,g} - \hat{p}_{rp,k-1,g} \leq u_{rp,k,g} * RU_g \quad \forall\, rp,k,g \quad (2l)$$

$$R_{rp,k-1,s}^{intra} + R0_s - R_{rp,k,s}^{intra} + W_k^k * D_{rp,k,s} + W_k^k * q_{rp,k,s} + W_k^k * p_{rp,k,s} - \mathrm{sp_{rp,k,s}} = 0 \quad \forall\, rp,k,s \quad (2m)$$

$$(R_{t-M,s}^{inter} - R_{t,s}^{inter}) + R0_{t=M,s} + \sum_{h_{t,rp,k}} (D_{rp,k,s} * W_k^k - p_{rp,k,j} * W_k^k + q_{rp,k,s} * W_k^k - sp_{rp,k,s}) = 0 \quad \forall\, t,s \quad (2n)$$

$$u_{rp,k,g}, y_{rp,k,g}, z_{rp,k,g} \in \{0-1\} \quad \forall\, rp,k,g \quad (2o)$$

Now that variables have been defined, the equations of this model are listed. Equation (2p), known as objective function, represents the total power system costs' that are desired to be minimized in [M€]. This constraint takes into account start-up costs, commitment costs, cost of energy production and of operation and maintenance, as well as the cost of power non served linked to the potential occurrence of this singularity. In this constraint, spillages in the hydro reservoirs are penalized with half of the energy's non served cost. The two variables linked to the startup decisions as well as the unit commitment (connection of the generation unit) have its associated cost ($C_j^{int}$and C_j^{st}) as previously explained. Complementary, the power production in [GW] is multiplied by the unitary cost in [M€/GWh] of the generating unit producing at each hour.

All non-served power, hydro spillages, binary commitment and startup and production variables are multiplied by the weight of each representative period and periods inside itself: $W_{rp}^{rp} * W_k^k$ measured in hours.

Equation (2q) represents the time slice power balance equation for the active nodes in the power system. The hourly demand $\mathrm{D_{rp,k,i}}$ of each period k and node i (included in a representative period rp) equals the positive power flows of input lines and the negative power flows outputs of each node; additionally it is necessary to sum the total production in [GW] of each generator connected to node i, which is noted by gi, and subtract the storage consumption taking into consideration the amplifier factor of the efficiency's inverse.

Equations (2r) and (2s) establish the maximum production in [GW] over the minimum production, known as P_g^- , of each thermal unit in relation with the binary commitment, startup and shutdown variables. This production range between the maximum and minimum generated power is noted with the difference of both up and down limits $(P_g^+ - P_g^-)$. An example of how binary variables work and how its differences in equations (2t) and (2u) have implications in the power plants' productions can be seen in Figure 3. Additionally, equation (2v) defines the binary nature of commitment, startup and shutdown variables: $\boldsymbol{u_{rp,k,g}}, \boldsymbol{y_{rp,k,g}}, \boldsymbol{z_{rp,k,g}}$.

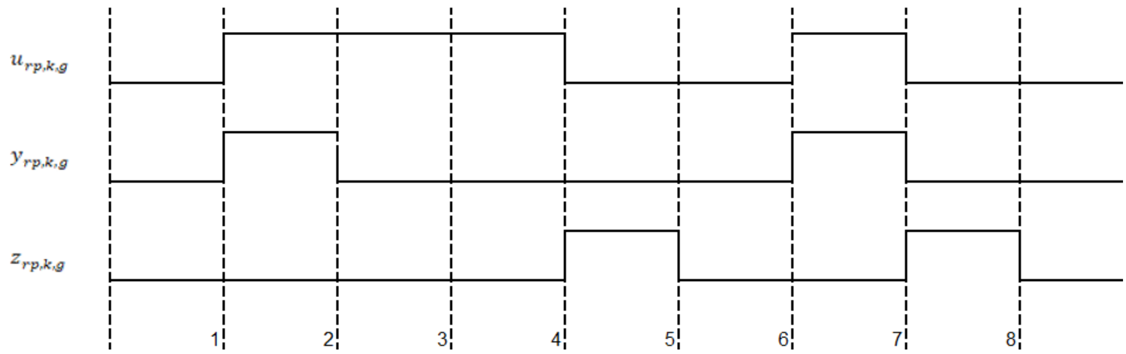


Figure 3. Equations (2w) and (2x): interactions between binary variables.

This Figure is representative since it also explains how equation (2y) works. Additionally, the generators production in [GW] is calculated as the minimum produced power if the unit is connected/commited ($u_{rp,k,g}$=1) in addition to the additional production over the minimum, listed as $\hat{p}_{rp,k,g}$ in the model formulation.

The last thermal equation is (2z) and it bounds the maximum power (over the power floor) that can be ramped up and down betweeen consecutive time slices (hours in the yearly model, periods inside a single representative week in the Chronological model or representative days in 7LRP model).

The constraints involving storage capabilities are (2aa) and (2bb). The first one applies to all models and balances the storage reserve in every hour or period inside its representative period $\boldsymbol{R}^{\boldsymbol{intra}}_{\boldsymbol{rp,k,s}}$ in [GWh] by ading the reserve of the previous period $\boldsymbol{R}^{\boldsymbol{intra}}_{\boldsymbol{rp,k-1,s}}$ and substracting the reserve dedicated to following periods and calculated in the period k in which equation (2cc) calculates the storage balance. In addition, the term $\boldsymbol{R0_s}$ is only summed in FullYear and Chrono models and represents an initial reserve in the intra-period of a whole year (Base Case) or a clustered week (Chrono model approach). If the modelled storage is Hydro, it will be necessary to sum the inflows for hydro storage $D_{rp,k,s}$ in [GWh] weighted with the parameter W^k_k representing the hourly weight for each period k. Logically, hydro units also comsume a variable amount of energy in [GW] modelled with $q_{rp,k,s}$, that must be taken into consideration in this constraint as a positive amount. Lastly, it is summed the storage production $p_{rp,k,s}$ also in [GW] and substracted the corresponding amount of spilled energy $sp_{rp,k,s}$ in the hydro storage case in [GWh].

The last equation (2dd) only applies to the 7LRP model, whom introduces the parameter M, a moving window in [h] with whom it is possible to to an inter storage balance -only in hours that are a multiple of this parameter M. In other words, with this equation it is possible to model not only storage inside a representative period but also to do so in-between representative days of a same week in 7LRP. As done in equation (2ee), a new variable $\boldsymbol{R}^{\boldsymbol{inter}}_{\boldsymbol{t,s}}$ is included in order to calculate the reserve at hour t and sum the previous period storage reserve $\boldsymbol{R}^{\boldsymbol{inter}}_{\boldsymbol{t-M,s}}$. There will also exist an intial inter period reserve, called $\boldsymbol{R0_{t=M,s}}$ from whom this equation departs from. The sum over the index $h_{t,rp,k}$, representing rhe relation among periods and representative periods rp, of the inflows and units consumptions', minus the spilled storage energy and unit production, finally balances all terms implicated in this restriction.

3. Base Case

This section includes a description of the benchmark (base case) which is useful for representing the real optimization problem. This benchmark really corresponds to the Full Year model, solved with unit commitment, so to say with a MIP approach. Section 3.a includes a summary of its results. Further comparisons begin with Section 4, in which, for every model, rMIP programming is studied over the real unit commitment MIP approach for computational time execution reasons.

a. Base Case Results

The resulting objective function takes a value of **673,77M€.** Computational time solution and number of variables are as follows in Table 2.

MODEL	FULL YEAR
APPROACH	MIP
Objective function [M€]	673.768
Execution time [s]	21605.111
Number of discrete variables	340704
Number of equations	672673

Table 2. Main output for full year MIP model

Of all the generation sources considered in first place, the final optimization results shed light on the real generation of each unit, which can be seen in Figure 4. This generation mix is just representative for further analysis and comparison between models, approaches and this base case. Figure 4 details the presence of nuclear and CCTG generation units as base of the power electic system with the renewable support of wind an solar energy. This renewable presence of energy sources is relevant since a further analysis of their penetration in the power system will be studied.

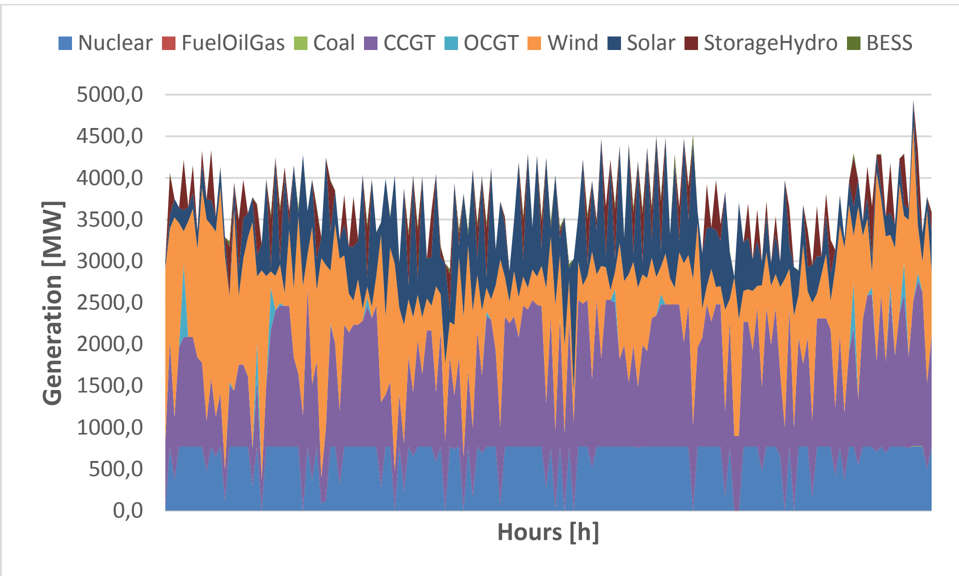


Figure 4. Generation mix for full year MIP model

Figure 4 informs that nuclear energy mainly serves as base of the generation mix scenario of a whole year of the operation of the power system. It produces its peak generation along many periods during the year and that is the reason why its generation appears horizontally cut at its maximum production. Since there are 4 CCTG units, its maximum generation is greater than the single nuclear unit previously mentioned. This type of technology is well used since the fuel cost is relatively low and has a variable cost much affordable than OCGT or Coal. This last two technologies are only taken by the optimization model when no other type of generation (including renewable sources) are able to cope with the demand.

Figure 4 also reveals a higher use of Solar energy in central hours and periods (hours that englobe light presence during a day and periods that correspond to spring and summer). On the contrary, Wind generation appears to have a more predictable profile except in the first periods of winter.

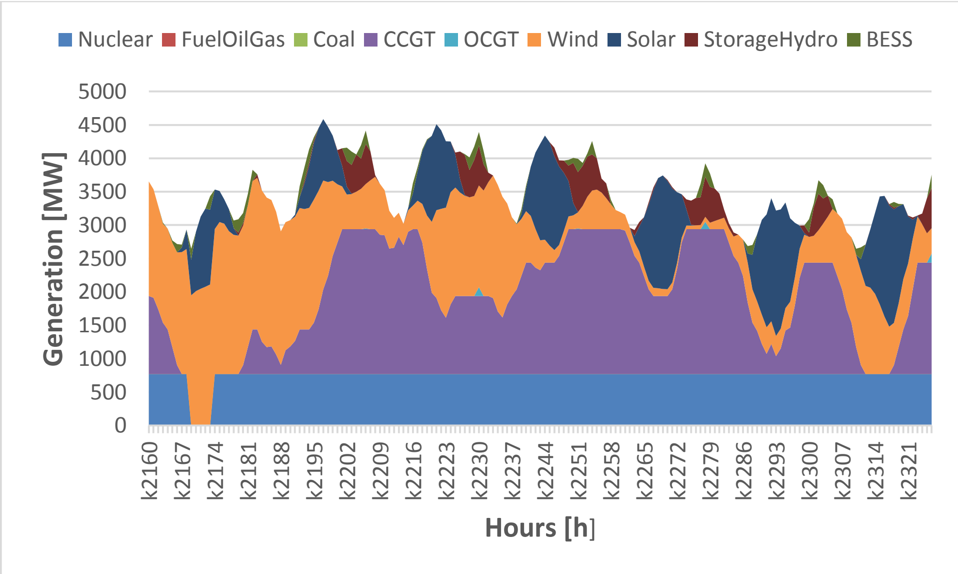


Figure 5. Detail of full year MIP model generation mix

Figure 5 gathers a representative week in favour of tracking each technology generation more effectively.

Another crucial point to be studied is the storage evolution-both batteries (BESS) and hydro units and for that reason Figure 6 and Figure 7, are included.

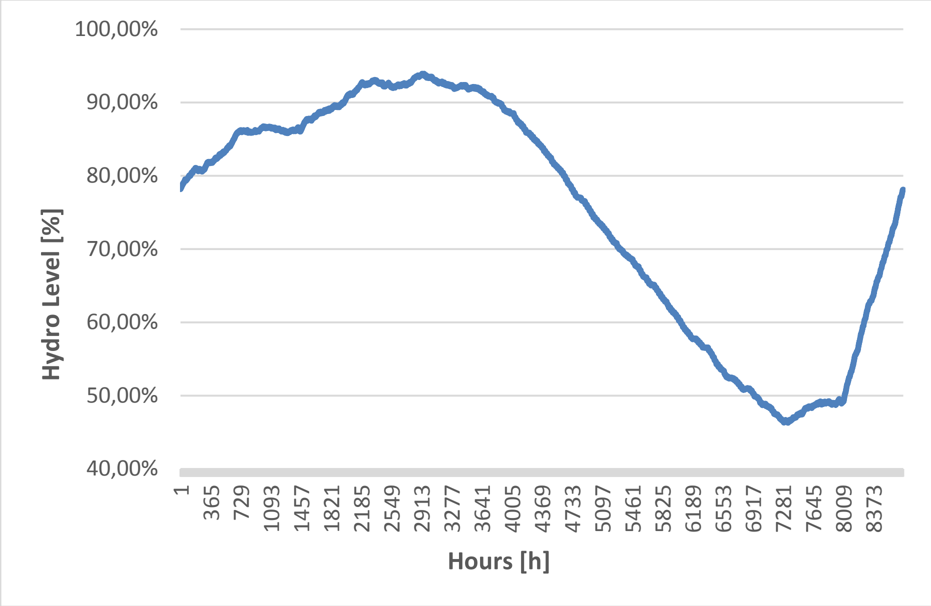


Figure 6. Hydro evolution for full year MIP model

As expected, the first months of the selected year usually increase hydro levels due to atmospheric conditions, reaching a peak of 94.02% in the hydro storage level. From this point onwards, warmer months tend to decrease hydrological reservoirs up until generally October, month in which Spain recovers the ability to boost hydro levels before a 47.44% off-peak storage hydro level.

In order to track the battery evolution of this annual model, a representative week is chosen as seen in Figure 7. Figure 7 shows the precision of a relaxed approach to the initial problem, although 164.a will broaden in the impact of a specific programming approach over another.

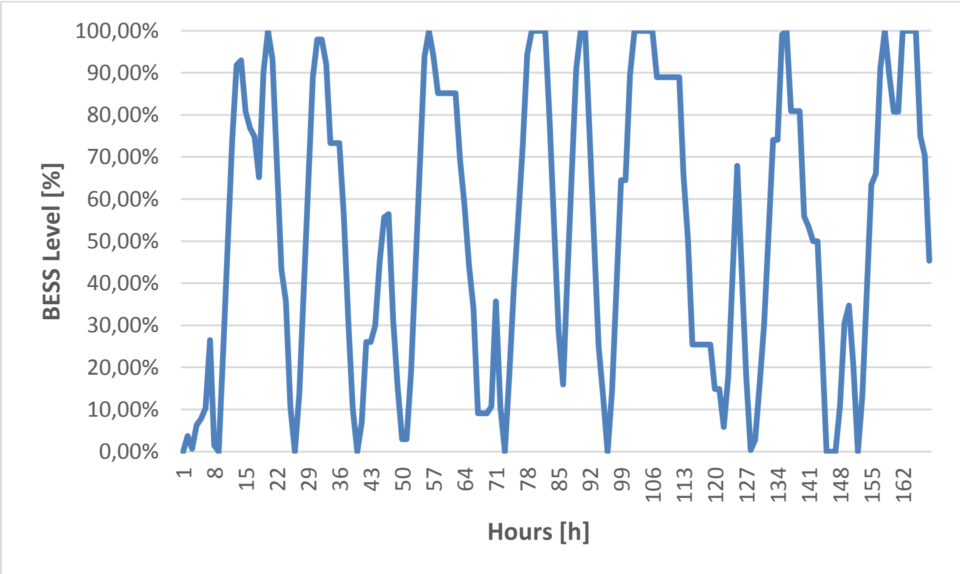


Figure 7. BESS evolution of a single week for MIP full year model

Figure 7 opposes with the hydro curve smoothness since BESS mainly serve for the storage of renewable generation. This is the reason behind a high variability of the storage level in low time slices. In fact, Figure 5 displays BESS levels above all productions; in other words, acts as a peak generation in the net hourly power balance.

The hydro reservoirs in Figure 6 confirm the hypothesis mentioned in 1.a of this document which is the monthly or even yearly work cycle. This fact contrasts with the battery duration of its charge/discharge cycle, generally around 4 hours.

4. Intra-Model Comparison

This section will be dedicated to analysis of the impact of the programming approach (rMIP or MIP) in the results of each of the models detailed previously.

Since the reduction of the total execution time is one of the main objectives of power system modelling and this can be obtained relaxing the initial unit commitment that MIP models imply, this different approach is considered. Even so we must maintain precision levels with this new approach. Only by maintaining precision levels, a reduction of computational time will make sense.

To examine these precision levels, we have included an estimation of the objective function error, listed in Table 7, as well as the consequent reduction of computational time employed by each rMIP programming compared to the real unit commitment of each model.

a. Full Year

Table 3 shows no sign of equation reduction or variable decrement since the objective of this section is to enlighten the results variability within the model itself and its relaxation of the unit commitment restriction.

MODEL	FULL YEAR	FULL YEAR
APPROACH	MIP	rMIP
Objective function [M€]	673.768	663.729
Execution time [s]	21605.111	37.977
Number of discrete variables	340704	340704
Number of equations	672673	672673

Table 3. Summary for Full Year model: MIP and rMIP

Restrictions concerning the hydrological level throughout the whole year impose the same performance for rMIP and MIP approaches so, as seen in Figure 8, they do not present relevant differences. Generally, the MIP model has slightly lower hydro levels than the rMIP, in other words: MIP makes a higher use of the available hydrological power than the relaxed alternative. Additionally, Figure 8 presents the equivalence of both programming approaches of the same annual optimization model, due to the lack of significant changes in the hydro storage levels within a year.

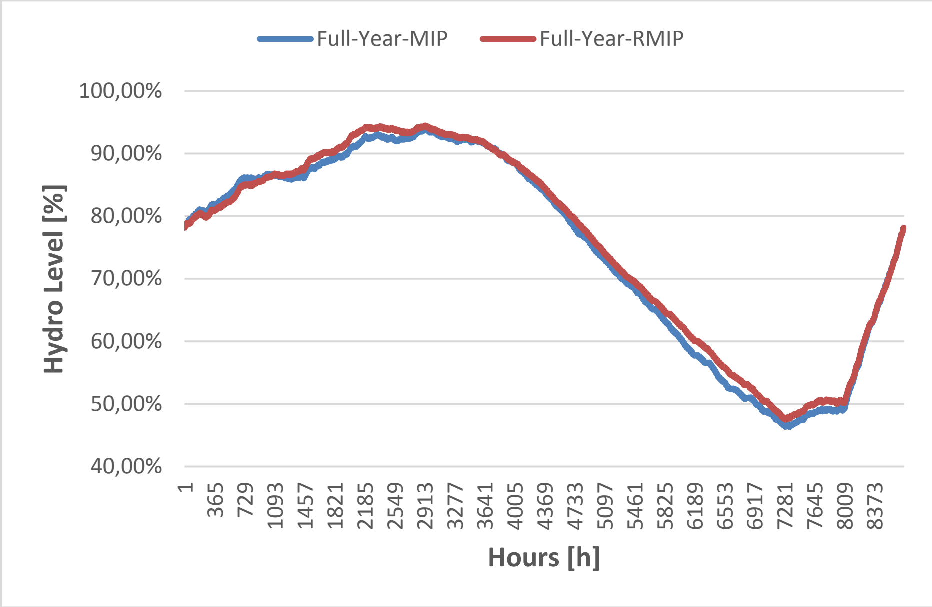


Figure 8. Hydro Evolution for Full Year Model

A deep look at the BESS Level for both approaches in the whole yearly scenario can be seen in Figure 9. This figure reflects how rMIP Full Year model loses details in the evolution of battery storage when comparing itself to the real MIP optimization problem. In spite of this fact, both models behave similarly, being the rMIP approach much quicker. For general and immediate results rMIP is a great alternative although the MIP is the authentic/real one.

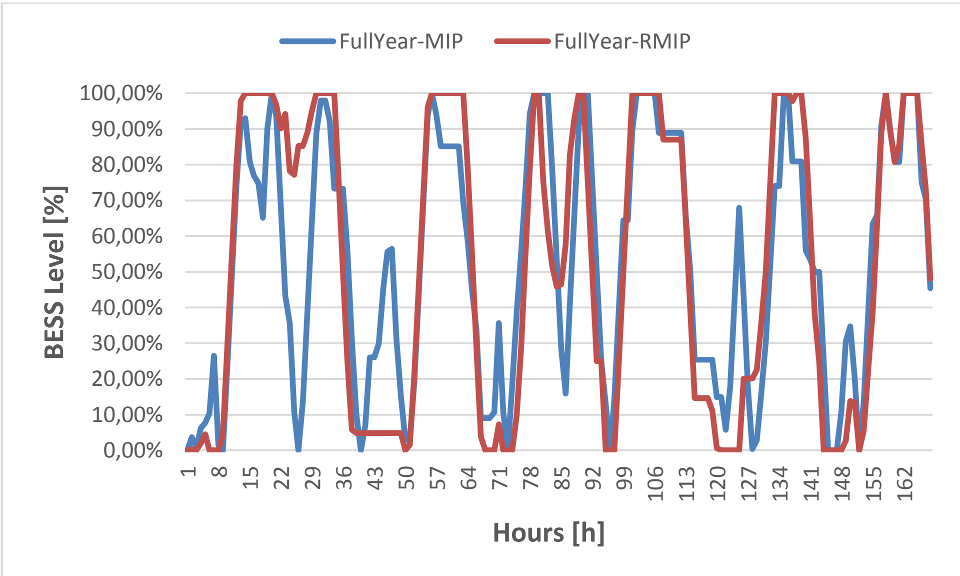


Figure 9. BESS Evolution Comparison of a single week for the Full Year Model

In order to quantify how much every rMIP approach differs from the original MIP model (precisely the variance seen in Figure 9), Table 8 is created. Although this table is included in the transversal intra model comparison in Section Transversal Intra Model Comparison5, a total BESS error of 9.12% is made in the Full Year Model.

b. Chrono

Stablishing as base of the comparison the Chrono MIP model, Table 4 is obtained:

MODEL	CHRONO	CHRONO
APPROACH	MIP	rMIP
Objective function [M€]	685.875	673,445
Execution time [s]	842.482	0.42
Number of discrete variables	6552	6552
Number of equations	12937	12937

Table 4. Summary for Chrono model: MIP and rMIP

In order to track the Hydrological levels through a year, a representative week is chosen in the Chrono model, as seen in Figure 10, which also includes the optimization deviation of an approach over another. Although in general terms hydro generation makes a null mistake, it can be seen that some hours have a peak fluctuation of +6%, outweighed by other hourly periods where deviations are below the real generation scenario. It can also be noticed a greater sharpness in the MIP programming due to unit commitment jumps in other generation units.

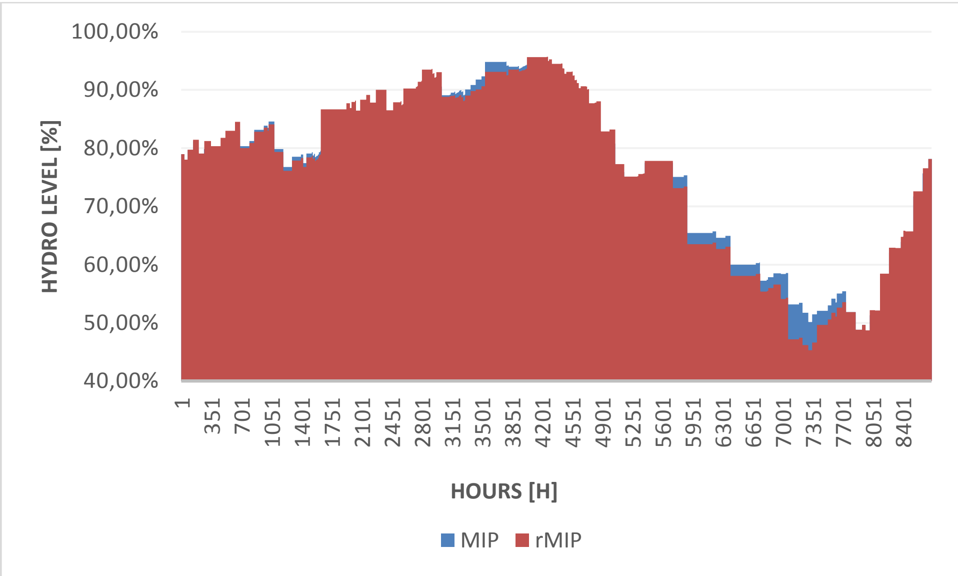


Figure 10. Hydro Evolution Comparison for the Chrono Model

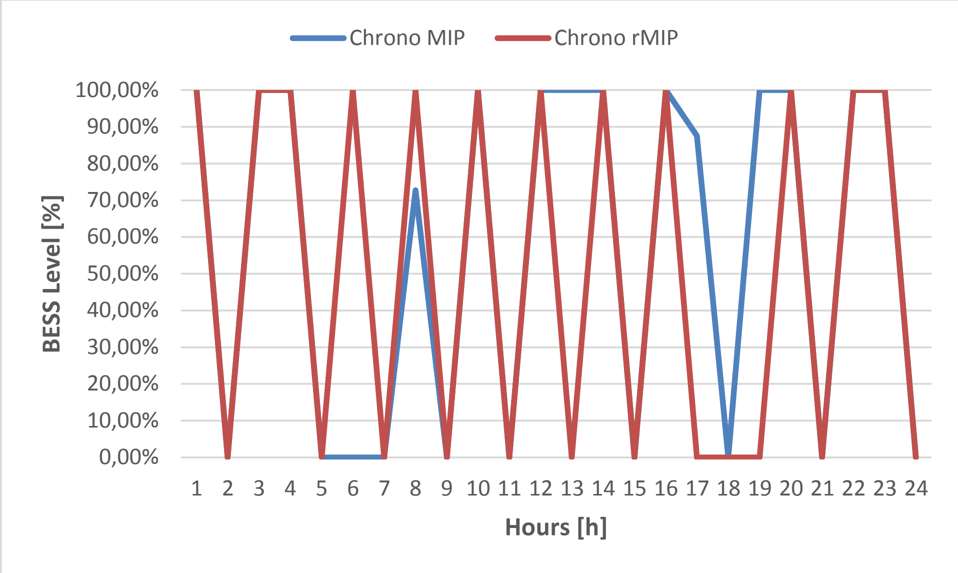


Figure 11. BESS Evolution Comparison of a single week for the Chrono Model

The trackability of battery levels is complex for them being truly changeable in a low time grid, in other words: within very few hours. Because of this fact Figure 11 only includes the fluctuation of storage levels in a chosen representative day. Selecting a wider time period would mean to increase operation jumps from 0 to 100% in the charge cycle or the other way around when batteries discharge. That being said, some pertinent differences are easily seen, that translate in a BESS error of 30.94%.

c. 7LRP kmeans

The only difference of 7LRP kmeans and 7LRP kmedoids models is the clusterization algorithm to aggregate data so in the end, energy production, storage evolution and hourly price changes behave alike. On this basis, the approach comparison of both cases is identical even though the results are numerically different so in order to avoid the repetition of contrasts, only one of them will be fully developed. Table 5 sums up the general optimization results of this case.

MODEL	7LRP kmeans	7LRP kmeans
APPROACH	MIP	rMIP
Objective function [M€]	655.430	644.900
Execution time [s]	47.099	0.462
Number of discrete variables	12124	12124
Number of equations	58954	58954

Table 5. Summary for 7LRP kmeans model: MIP and rMIP

Figure 12 represents the BESS performance for both approaches. It can easily be seen that the MIP model is more detailed and complex because it includes a graduated progression of the battery storage level. On the contrary, rMIP approach characterizes for being smoother although general behaviour of both of them is equivalent.

An interesting point to be distinguished is that MIP could have overlearned the proposed generation mix so if this is changed, the rMIP is expected to behave better since it can be generalized to other power systems. In other words, there should exist a good compromise between the results accuracy and the potential generalization to other generation scenarios.

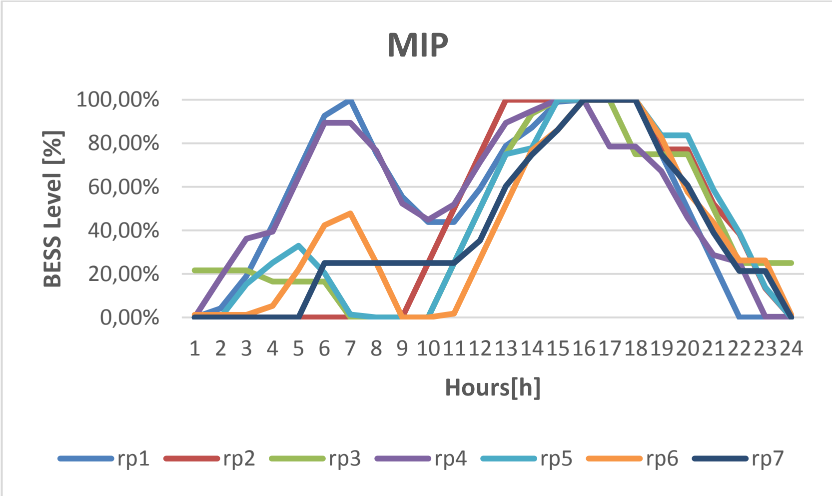
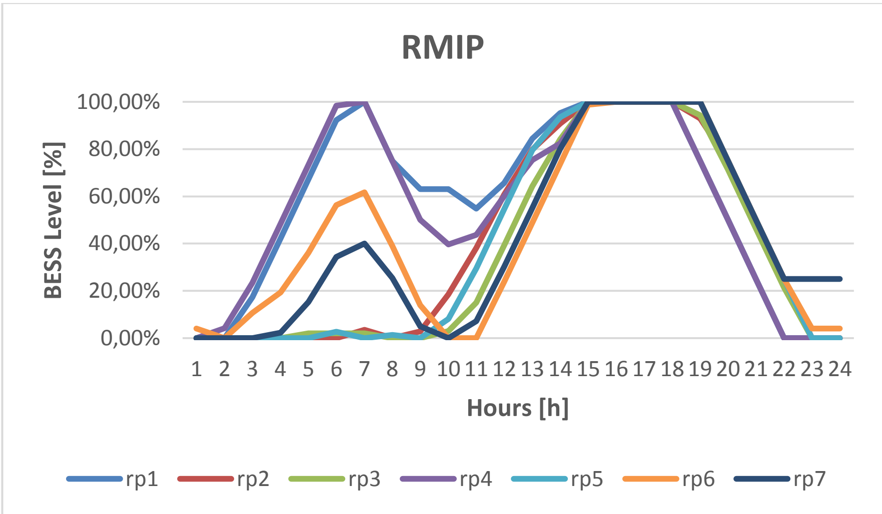


Figure 12. BESS Evolution for 7LRP kmeans: MIP and rMIP.



d. 7LRP kmedoids

As said in Section 4.c, 7LRP kmeans and 7LRP kmedoids models are equivalent so, in order not to repeat redundant information and programming approach contrasts, in addition to Table 6, which includes a performance summary of this model and clusterization algorithm, the evolution of hydro levels are displayed in Figure 13.

MODEL	7LRP kmedoids	7LRP kmedoids
APPROACH	MIP	rMIP
Objective function [M€]	649.784	639.148
Execution time [s]	54.433	0.524
Number of discrete variables	12873	12873
Number of equations	59136	59136

Table 6. Summary for 7LRP kmedoids model: MIP and rMIP

It can be easily seen that this the use of a relaxed approach displays better the levels throughout whom Hydro units go through the whole year in comparison to the same relaxation in the Chrono model. This can be noticed since the dispatch between one programming approach over the real MIP one is greater in Chrono model than in this case.

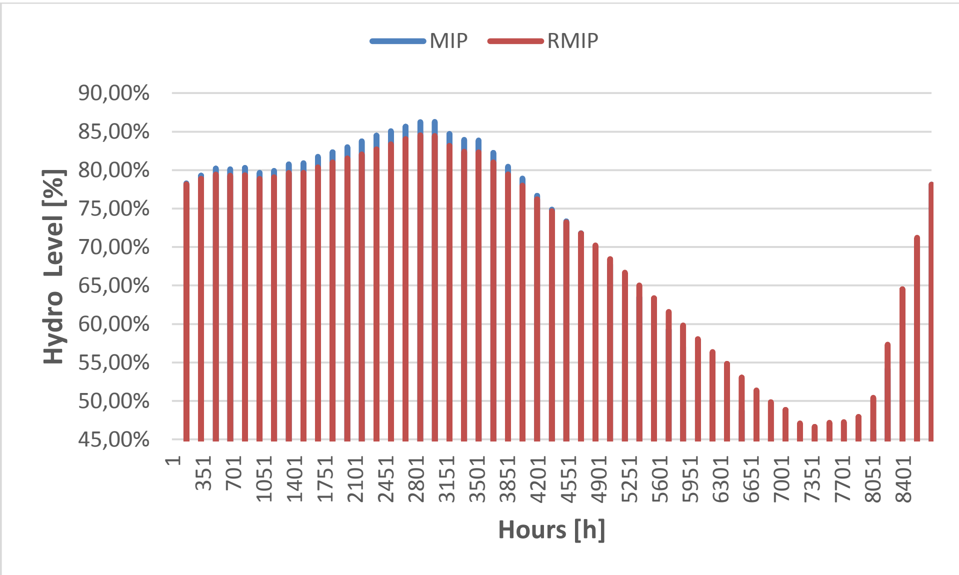


Figure 13. Hydro Evolution Comparison for the 7LRP kmedoids Model

5. Transversal Intra Model Comparison

	FULL YEAR	CHRONO	7LRP kmeans	7LRP kmedoids
Objective Function Error [%]	-1.49	-1.81	-1.61	-1.63
CPU Time Ratio	568.9	2005.9	102.0	103.9

Table 7. Comparison of Objective Function Error and CPU time ratio.

Table 7 resumes the objective function error in which each model incurs in case of selecting a relaxed programming approach over a unit commitment compromise, as well as a measurement of the reduction of CPU time. It is noted that the error in which every model incurs is negative since the MIP approach is always more expensive than the relaxed alternate model.

According only to this, some general conclusions can be drawn. Although in absolute terms, the Full Year model takes longer to be digested, a considerable reduction of time is seen between approaches, even a higher reduction than with smaller models. It can be seen that the rMIP approach to this model has many hours considered outliers since hours represent independent entities that can be "thrown away" from the relaxation of the generation constraint. On the contrary, Chrono and 7LRP models lack these outliers since a previous clusterization treats hours in groups -with a specific weight in Chrono model and representative days in 7LRP. The inclusion of clusters disables the probability of getting outlier groups or clusters, fact that can be seen on Figure 14.

At the same time, both clusterization algorithms in the 7LRP model happen to be comparable since they represent the same optimization problem solution. They both err in a nearly identical objective function deviation and reduce the employed time in the same order of magnitude. Lastly, the Chronological model serves as a good compromise between its-a bit higher-objective function disparity and an incredible drop in CPU time.

According to the objective function error and its fluctuation not being sustainable, it has been chosen to evaluate the error in hourly price, as shown in Figure 14.

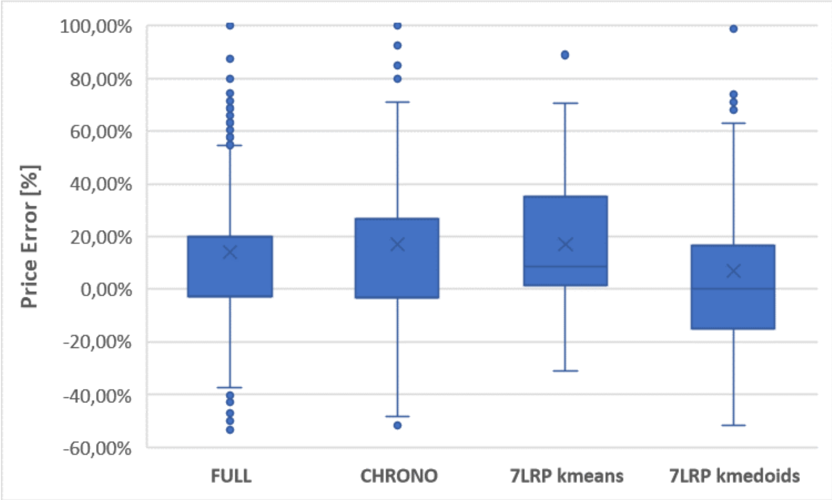


Figure 14. Price Error [%] in each model

I do consider relevant to link Figure 14 with future Table 8, due to the existence of two aspects that confirm its relationship: the fluctuation of the hourly price and the mean error in price. When looking at the price deviation, Figure 14 displays that, in general terms, 7LRP kmeans model has the highest variability and error since its root mean square error (RMSE) is the highest. RMSE is now useful for the reason of having positive and negative variations towards the role MIP model. For instance, the Full Year model, carries a smaller fluctuation in the hourly price error (it has the smallest RMSE) but entails a substantial number of outliers that correspond to a mistake in the generation technology that truly generates at all those specific outlier hours. That is the reason behind a general higher absolute error in all generation units in Table 8. To sum up, the variability in the price error is proportional to the value of the RMSE while the mean price error is directly proportional to the objective function error. In addition to the objective function value other variables must be studied to come up with a precise approach to the total. Let us now take a closer look to the production mix of each scenario. The total generated energy by each technology does vary as shown in Table 8.

	FULL YEAR	CHRONO	7LRP KMEANS	7LRP KMEDOIDS
Nuclear	-0,48%	-1,15%	0,00%	-0,10%
Fuel Oil Gas	0,00%	0,00%	0,00%	0,00%
BESS	-9,12%	-30,94%	4,00%	-13,52%
Wind	0,40%	0,54%	0,00%	0,00%
Solar	0,11%	2,00%	0,00%	0,12%
Coal	-44,36%	-99,45%	-100,00%	-17,77%
CCGT	1,79%	2,12%	1,68%	2,11%
OCGT	-39,62%	-24,67%	-63,68%	-58,76%
Hydro	0,00%	0,00%	0,00%	-0,20%
Total	-0.11%	-0.10%	0.04%	-0.19%
RMSE	36.24%	114.66%	140.74%	39.56%

Table 8. Summary of errors in a rMIP approach by technology type.

Table 8 also includes some expected results, such as a complicated modelling of coal as a generation source since this type of generation is only used in few days of the annual perspective. 7LRP model fails in modelling this generation since no representative days include isolated hours in which coal is used. Another important point to be distinguished is the treatment of storage units, in which hydro and batteries are included. Hydro generation is very well modelled by every relaxed approach although batteries have a larger error deviation in which they incur. On the other hand, BESS possess a wide variety of error: while they behave well 7LRP, Chrono rMIP fails more due to the formulation of the model itself: the weight or number of periods, affects the mean error with that coefficient so if the battery has a different storage level, it will be affected proportionally.

In general terms, the transversal intra model comparison throws interesting conclusions:

- The rMIP approach results always in a cheaper objective function value and differs from the true MIP model in an error bounded between 1 and 2% in absolute terms.
- The Full Year model behaves better when running the relaxed optimization problem than the Chrono and 7 Linked Representative Periods models (makes a lower mistake).
- The **CPU time** employed **in rMIP models is much lower** than the MIP one. Chrono model wins in time reduction, followed by Full Year Base Case (explained by the computational burden of this one) and 7LRP lastly.
- **The Chronological model serves as a good compromise with its objective function error and CPU time reduction if the optimization problem is multiobjective (lower CPU time drop and objective function disparity).**

According to price deviation and generation technologies, this section concludes:

- The **price fluctuation** is ordered in decrescent order in this way: FullYear<Chrono<7LRP. Despite this, outlier hours present a higher presence in FullYear>Chrono>7LRP.
- **Price fluctuation** is directly proportional to the **RMSE** that each model errs in.
- **Mean error** is higher in Chrono than in 7LRP and Full Year models.
- **Coal** technology is badly modelled since this generation type is only used in few days of the annual perspective and clusterization algorithms do not consider it as yearly representative.
- **Hydrological reservoirs** are well modelled by every approach. **BESS** on the contrary behave better in 7LRP>FullYear>Chrono. This is useful because 7LRP will be a better alternative when modelling batteries.

6. Transversal Inter Model MIP Comparison

	CHRONO	7LRP KMEANS	7LRP KMEDOIDS
	MIP	MIP	MIP
Objective function [M€]	1.80%	-2,72%	-3,56%
Execution CPU Time Ratio	25.646	458,717	396,912

Table 9. Summary of Model comparison

In general terms, the model which best behaves according to the optimization results seen in Table 9 is the Chrono model. As another extra general conclusion, it can be extracted that the kmeans algorithm leads to a more reliable and faster clusterization method than kmedoids, having the same equations and variables load. It is also relevant to denote that the variability with the kmedoids algorithm is greater as the representative periods correspond to real days.

Figure 15 reflects the price evolution for all models as a measurement of the Objective Function Error and the variability of generation technologies that fix hourly prices.

As seen in Figure 15, the Chrono MIP model catches the exact price fluctuation the Full Year MIP model has, which is the reason behind a good behaviour in error terms. 7LRP models are expected to work as seen in Figure 15 because the clusterization algorithm itself does not consider representative hours with price fluctuations over a certain value. This opposes with the Chrono model, able to restrain these hours, maintaining an accurate generation production if compared with the base case.

It can also be noticed the same tendency through the MIPs and rMIPs approaches: generally, the price deviation is wider in MIP, being more reduced in relaxed terms.

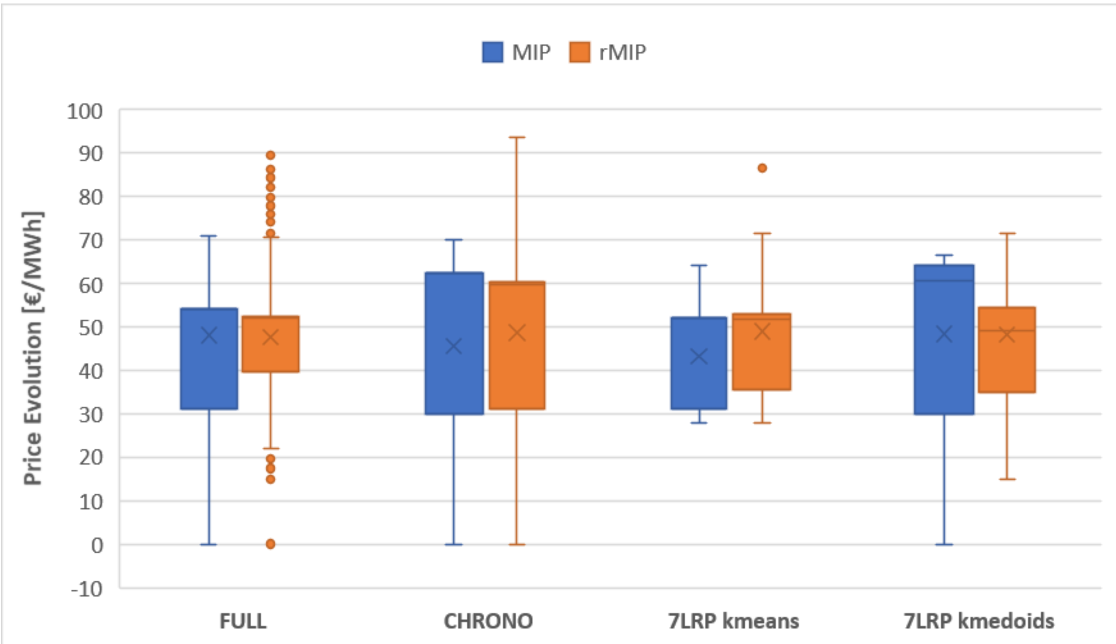


Figure 15. Price evolution Inter-Model

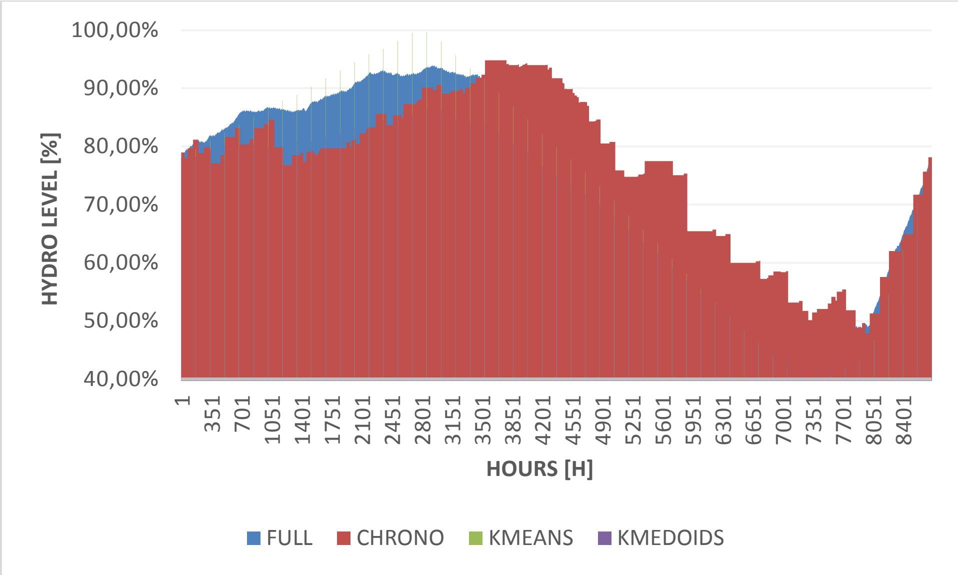


Figure 16. Hydro levels Inter-Model.

Hydro levels are registered in Figure 16 in order to illustrate the difference between the different models' behaviour and differences. Kmedoids seems to be the only model that lacks in modelling Hydro Technology perfectly, fact resumed in Table 10.

	CHRONO	7LRP KMEANS	7LRP KMEDOIDS
Nuclear	-4.05%	2,34%	3,74%
Fuel Oil Gas	0,00%	0,00%	0,00%
BESS	-75.07%	-12,10%	13,82%
Wind	0,26%	0,55%	0,35%
Solar	-2.04%	0,29%	7,34%
Coal	-26.14%	-77,68%	-24,80%
CCGT	3.41%	0,15%	-6,10%
OCGT	37,18%	21,40%	100,70%
Hydro	0,00%	0,00%	-3,60%
Total	-0.94%	-0,15%	-0,20%

Table 10. Error in Inter-Model Generation Comparison

Table 10 displays the optimization error of MIP model approaches in comparison with the base case. The total generation error is equivalent in all models although 7LRP kmeans acts with better accuracy than Chrono model and the same but with kmeans algorithm. Although Coal Generation is badly optimized-because this type of generation is only being considered during very few days in the year- in operation regimes, this mismatch is not relevant. On the contrary, if a future investment analysis is proceeded, those 'coal' extreme days would make a difference.

It is important to underline that not only the model is better by having generally lower errors in every technology but making low mistakes in more frequent and used technologies. Since Coal, BESS and combined cycles are generation technologies used at peak consumption, they are not entirely representative. Nevertheless, nuclear energy represents part of the generation base and has a higher presence in the generation scenario (and consequently more weight). According to this, Table 10 places 7LRP better than the chronological model because it has a smaller fluctuation in nuclear generation.

This section concludes some remarkable points:

- According to the **error** made in the objective function value: the **Chronological model is the best**, followed by 7LRP with a kmeans clusterization.
- Despite the previous point the higher CPU time reduction is made in **7LRP kmeans and kmedoids (they are quicker but less precise). Chrono model is more accurate but slower.**
- **Between kmeans and kmedoids**, the first one behaves better in both objectives (objective function value and CPU time). **Kmeans algorithm leads to better optimization results.**

According to hourly price deviation:

- Chrono catches the real price fluctuation, 7LRP models do not since the time structure is very different and Chrono and Full are structurally alike.
- **Price deviation is wider in MIP**, being more reduced in relaxed terms.

According to storage technologies:

- Kmedoids is the only model that lacks in modelling Hydro Technology perfectly, incurring in a 3.6% error. **They all model well the hydro storage.**
- **BESS are always better modelled by 7LRP** (both algorithm approaches) because they present a time structure based in hours and representative days, period in which batteries work, and unlike the Chrono method which has longer time slices than representative days.

7. RES Sensitivity Analysis

This section is dedicated to examining the impact of a changeable Renewable Energy Generation Sources penetration to the model. Both Wind and Solar RES in the Base Case represent a 35% of the generation mix considered. The methodology proposed in Section 7.a is considering a full wind 35% generation whether as Section 7.b does the same with Solar energy. Before examining both cases separately, an initial analysis is conveyed with interesting general results.

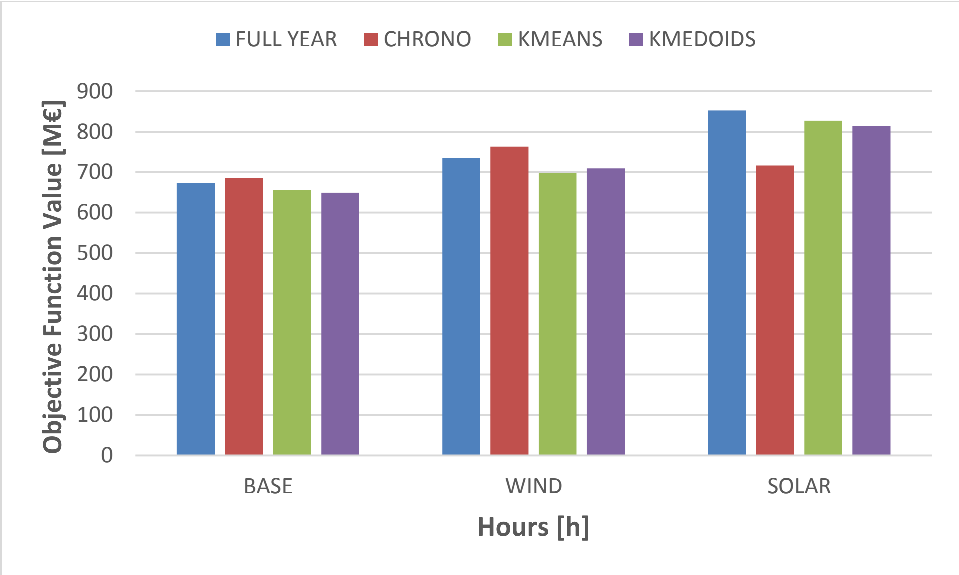


Figure 17. Objective Function for the three RES cases.

Firstly, let us now comment general expected results obtained towards the value of the objective function. As it may be seen in Figure 17, the higher the RES penetration is, the higher the total costs are. A generation mix as one in the base case with both wind and solar energy, is, of the cases considered in this section the 'optimal' since a change in producing null Solar energy or doing alike with the wind production leads to a higher objective function value. Since the wind production is unitarily cheaper than the Solar generation, for all models, wind is more expensive than the Base Case, but cheaper than having a renewable scenario with only solar production.

Secondly, let us now compare these results in-between the models used is this thesis. Again, since the Full year model is the 'real' one, Table 11 resumes the objective function error each model has in comparison with yearly data.

	Chrono	7LRP kmeans	7LRP kmedoids
Base Case	1.80%	-2.72%	-3.56%
All Wind	3.70%	-5.19%	-3.55%
All Solar	-15.96%	-2.93%	-4.49%
RMSE	**16.48%**	**6.55%**	**6.74%**

Table 11. Model and Case objective function Error

Curiously, modifying renewable penetration affects in which model is the best to optimize this problem. In other words, now we do not have a better optimization model and it will be necessary to analyse which renewable generation mix we have. Table 12 orders all models in crescent error to determine the better for each mix. It is also important to underline the adaptability each model has towards a change in the penetration of every type of technology, that is why in Table 11 it has been included the RMSE of each model as a representation of how adaptable each optimization model is. The 7LRP kmeans is the best way of reducing costs if the renewable penetration is variable while on the contrary and curiously Chrono model is the worse, maybe because it has already overlearned the Base Case data, which disables its adaptability.

Generation mix	1st	2nd	3rd
Base Case	Chrono	Kmeans	Kmedoids
All Wind	Kmedoids	Kmeans	Chrono
All Solar	Kmeans	Kmedoids	Chrono

Table 12. (Best) Ordered models by RES generation mix

There is another funny result obtained in this section, which is the presence of renewable and not used generation; in other words, spilled, since it is not cost-efficient. Logically, in Figure 18 the Solar generation waste is higher due to the limited capacity of the batteries included in the system, that are unable of store higher power levels in central hours of each day, where there is still daylight to benefit from. This does not happen equally with wind production due to a higher variability and unexpected changes in its behaviour.

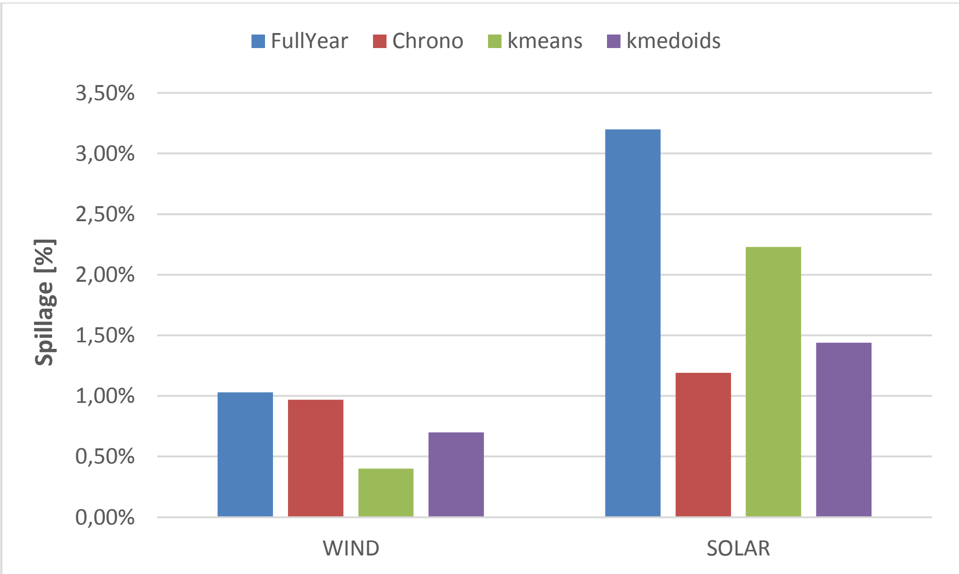


Figure 18. RES Spillage [%]

Also, the presence of renewable spillage has a direct effect over the total power system costs, having higher objective function values as higher the spillage is.

Regarding model's performance, it can be observed that the yearly model is the one spilling more energy because it is able of realising an hourly difference that other clustered models have lost in the process of merging time slices.

General conclusions are included in the following points:

- **The Objective Function Value is directly proportional to the presence of RES**: having more wind or solar leads to more expensive general results because unitarily, solar energy is more expensive than wind production.
- Having a generation mix with **high penetration of renewable sources** leads to a better modelling on the part of **7LRP methodology** (in comparison with the Chronological model)
- RES Spillage is higher in a scenario where renewable production only depends on solar generation (than in the same one but depending on wind).
  Full Year model is the least optimal since it can see the spilled energy through all hours in the yearly perspective.

a. Wind

In this section we want to study the impact of having a total renewable generation just depending on wind production. A summary of the optimization results is gathered in Table 13.

	FULL YEAR	CHRONO	7LRP kmeans	7LRP kmedoids
Objective function [M€]	735,971	763,232	709,843	709,843
Objective Function Value Difference [%]	9,23%	11,28%	8,30%	9,24%
Number of discrete variables	340704	6552	6552	6552
Number of equations	672673	12937	12873	12873

Table 13. Summary of results: RES Wind production.

Again, Table 13 shows a sign of each model's adaptability towards changes in renewable production through the objective function value. This confirms that 7LRP kmeans is the option that best behaves if a there is a changeable RES penetration, being Chrono the worst model to adapt towards changes.

Taking a deeper look at the model's and technologies generation in comparison with the yearly perspective in Table 14, the main production differences that can be extracted are the coal, combined cycles and BESS generation. And the reason behind these technologies having their productions changed is, for coal and combined cycles (CCGT and OCGT) the lack of a considerable amount of these technologies' generation in the base yearly case; while, for the batteries difference it is the smaller duration of the clusters in 7LRP whom justifies better results in these models in comparison with the Chrono one.

	CHRONO	7LRP KMEANS	7LRP KMEDOIDS
Nuclear	-5,64%	4,96%	4,47%
Fuel Oil Gas	0,00%	0,00%	0,00%
BESS	-69,53%	-26,24%	8,28%
Wind	-0,69%	1,53%	0,70%
Solar	0,00%	0,00%	0,00%
Coal	-54,13%	-100,00%	-33,84%
CCGT	4,53%	2,44%	-0,95%
OCGT	481,23%	-100,00%	-100,00%
Hydro	0,00%	0,00%	-3,81%
Total	-0,87%	-0,33%	-0,26%
RMSE	489,29%	143,95%	106,06%

Table 14. Inter Model Comparison: RES Wind Production.

The better BESS treatment from the 7LRP model (in both clusterization algorithms) is illustrated in Figure 19, whom takes BESS levels' first representative day of 7LRP kmeans model and compares itself with the corresponding same day for the remaining optimization models. The visual comparison of the yearly model is very similar to the 7LRP whether the Chrono model fails in depicting intermediate levels, fact that translates into a worse response.

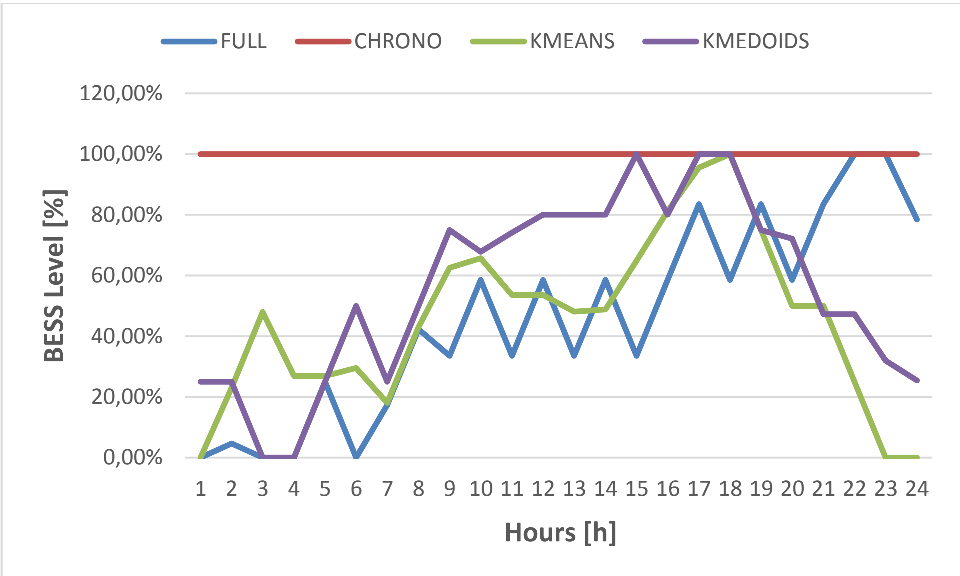


Figure 19. Inter Model Comparison: BESS Level

Figure 20 is included in order to track the evolution of the 7LRP model with a kmeans clusterization algorithm. It can be observed that a 35% wind presence has a great impact on the batteries' evolution and use. Moreover, the use is 'optimal' since the BESS do not have long periods in which they saturate at its maximum but charge and discharge gradually through the representative period (day).

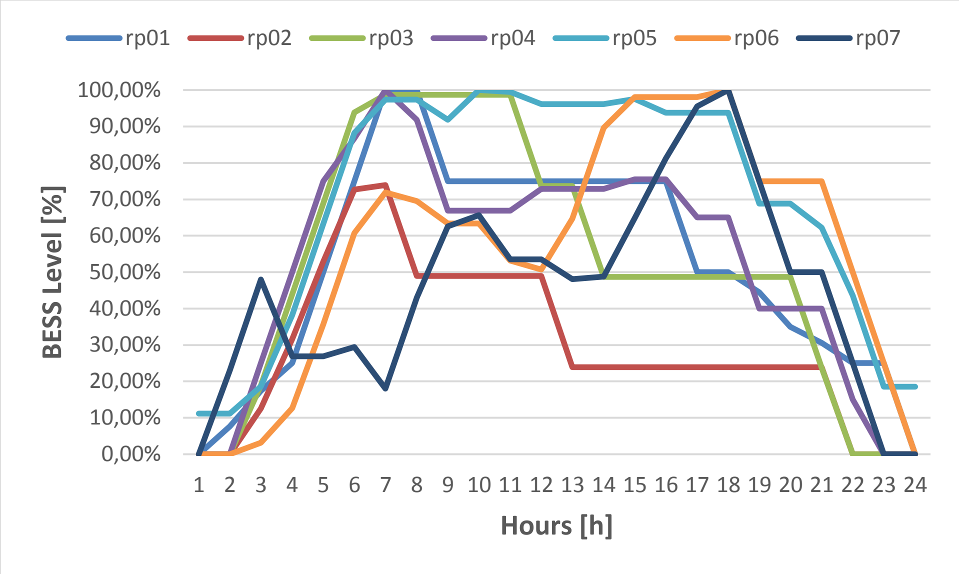


Figure 20. BESS Evolution: 7LRP kmeans RES Wind production.

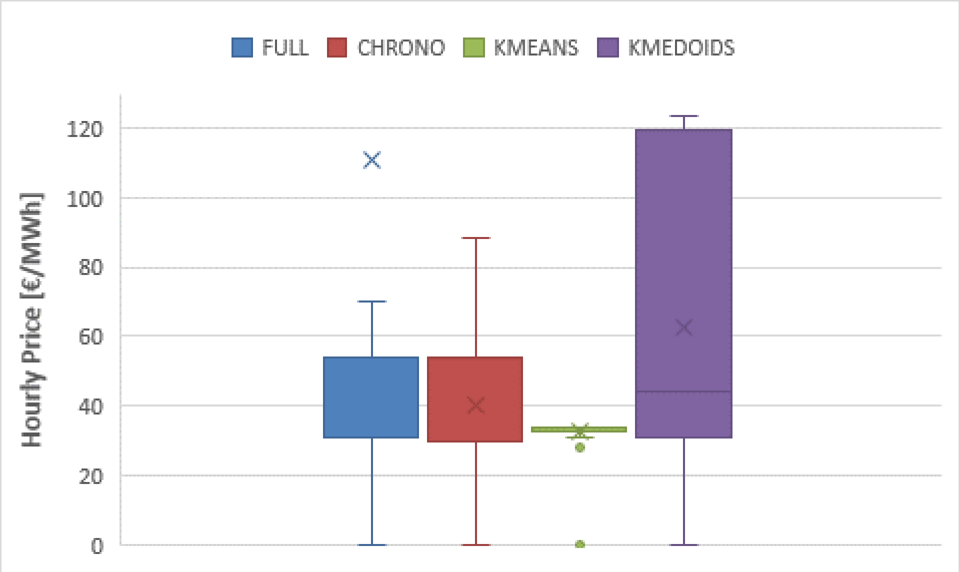


Figure 21. Hourly price evolution: RES Wind Production.

Conclusions on the wind scenario are gathered hereunder:

- **7LRP kmeans** is the model that **best behaves** if a there is a **changeable RES penetration.** On the contrary, **Chrono is the worst model** to adapt towards this kind of changes.
- BESS' treatment from part of 7LRP model is due to an equivalent resemblance of this approximation towards the Full Year model. On the contrary, the changeable BESS level of the Chronological approximation from 0 to 100% is not able of studying smaller BESS differences.
- The main **production** technologies **differences** derive form the lowest power generation from **coal, combined cycles and BESS discharges**.

- A **35% wind presence** has a remarkable **impact on the batteries' evolution and use**. This use is 'optimal' since the BESS do not have long periods in which they saturate at its maximum but charge and discharge gradually through the representative period (day).

b. Solar

In this section the objective is to estimate the impact of having a total renewable generation just depending on solar production. A summary of the optimization results is resumed in Table 15. This table includes an 'objective function difference' subsection that tries to estimate the impact of this solar mix in comparison with the initial generation mix proposed-with 35% RES presence. Having the same order of magnitude in the yearly and representative periods approaches, leads to a general better tracking from 7LRP model than the Chrono one, similarly as happened in Section 7.a.

	FULL YEAR	CHRONO	7LRP kmeans	7LRP kmedoids
Objective function [M€]	852,409	716,352	827,406	814,120
Objective Function Difference [%]	26,51%	4,44%	26,24%	25,29%
Number of discrete variables	340704	6552	6552	6552
Number of equations	672673	12937	12873	12873

Table 15. Summary of results: RES Solar production

To confirm this intuition, the hourly price each model goes through (determined by the last generation technology producing power) is illustrated in Figure 22. Naturally, the price variability in the 7LRP model resembles the yearly one, while the Chronological clusterization fails in modelling the hourly production well.

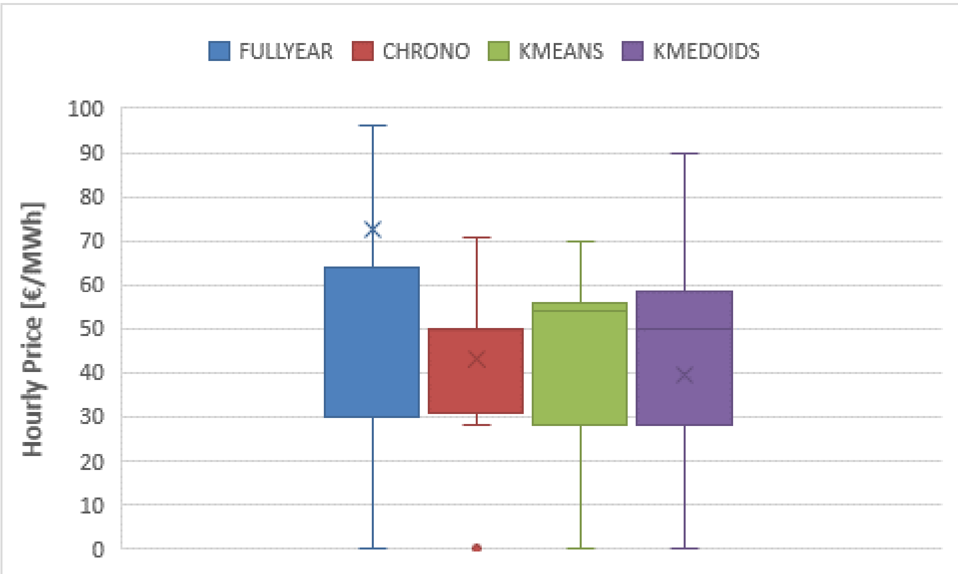


Figure 22. Hourly price evolution: RES Solar Production

Figure 22 shows that the Full Year model is affected by the outlier hours in which energy is produced by a much more expensive technology in isolated hours and days, that is why the mean hourly price of this method is much higher than with other studies.

Now again, the most affected generation technologies are the smallest ones (Coal, CCGT and OCGT) as well as the batteries use. There has been included a point in which the RMSE is calculated to evaluate the total production error of each model in comparison with the Full Year. This calculation approves the better 7LRP modelling.

	CHRONO	7LRP KMEANS	7LRP KMEDOIDS
Nuclear	14,83%	1,15%	-1,50%
Fuel Oil Gas	0,00%	0,00%	0,00%
BESS	-86,73%	-3,10%	13,33%
Wind	0,00%	0,00%	0,00%
Solar	4,99%	3,00%	5,37%
Coal	-73,50%	19,76%	-31,66%
CCGT	-2,13%	-4,85%	0,35%
OCGT	47,92%	-50,12%	-40,50%
Hydro	0,00%	0,00%	-3,54%
Total	-1,27%	-0,04%	-0,17%
RMSE	124,39%	54,28%	53,51%

Table 16. Inter-Model Comparison: RES Solar Production.

Figure 23 pictures one of the most affected technologies evolution, in this case, the BESS levels, throughout the seven linked representative periods. It can be seen that a relatively high solar presence leads to a rather predictable batteries use in every representative period (day) of the year. This fact could be used in computational benefit choosing a smaller number of representative periods in both kmeans and kmedoids algorithms.

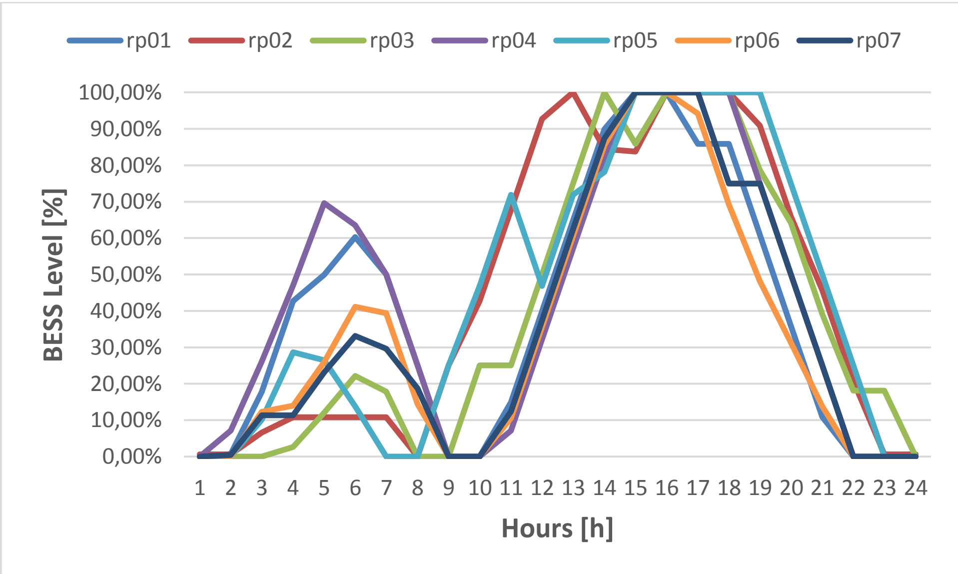


Figure 23. BESS Evolution: 7LRP kmeans RES Solar production.

Conclusions on Solar scenario generation include:

- **Chrono** model is the one that least varies in comparison with its complementary in the base case and logically the least accurate in comparison with the hourly model because it **is more rigid and reticent to changes in its generation mix**.
- **The price variability and smaller errors** are gathered by the **7LRP** methodology because:
- **BESS storage is better modelled** with a flexible LRP approach and because total **RMSE is lower** than with a chronological study.
- As noted previously, **batteries saturate** in central hours of the day leading to higher value of the optimization problem.

In order to keep a wider look at the carried sensitivity analysis just carried and determine all implications towards hourly price change, Figure 24 is resumed:

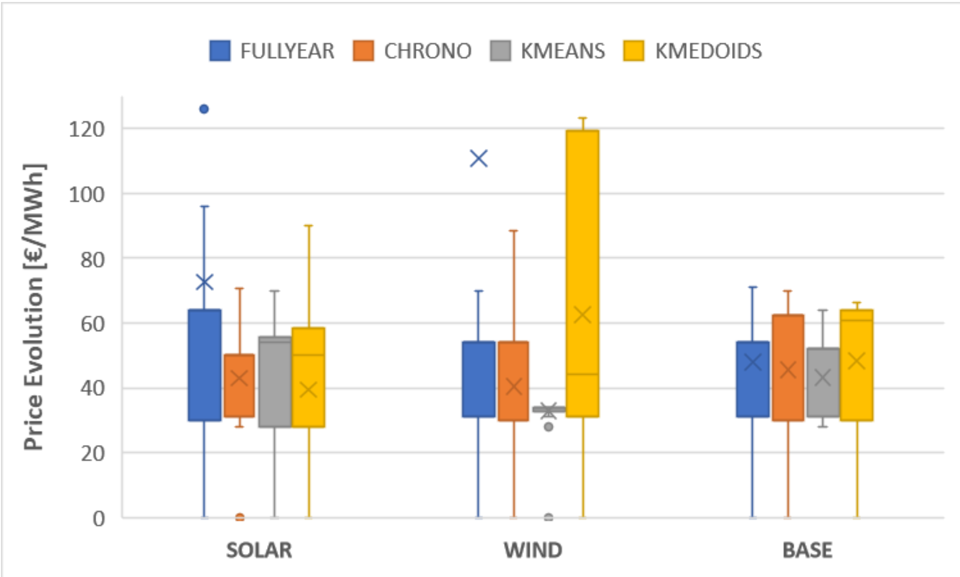
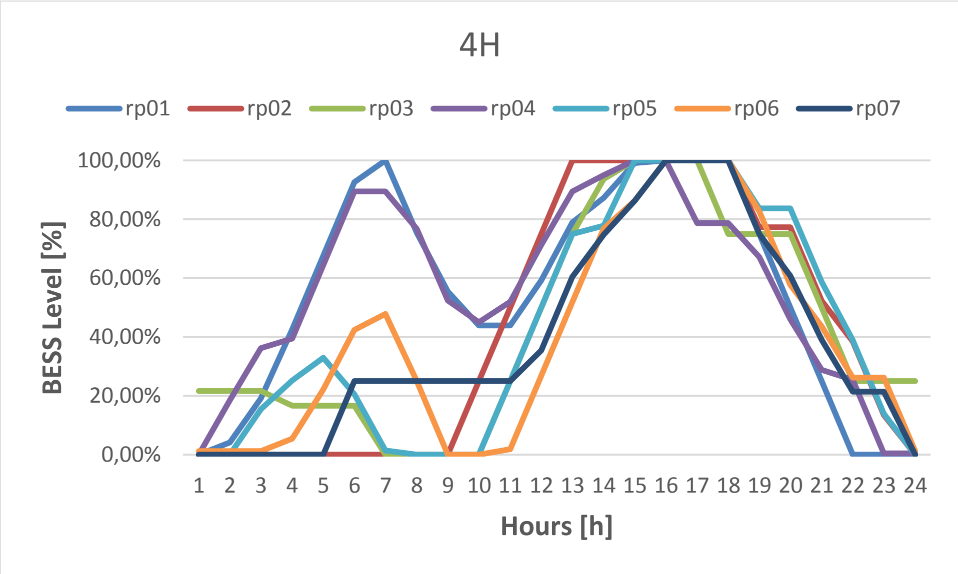


Figure 24. Transversal RES Analysis.

8. Transversal BESS Sensitivity Analysis

This section will be dedicated to the analysis of the change in the duration of the charge and discharge batteries cycle. This will be relevant since not all models behave equally, so the goal of this analysis is to study the evolution of all models based on the number of hours their duration cycles includes.

To see an example of how every model behaves (they all behave similarly) when enlarging its work cycle, 7LRP kmeans batteries level is more tractable and it is shown in Figure 25 and Figure 26.



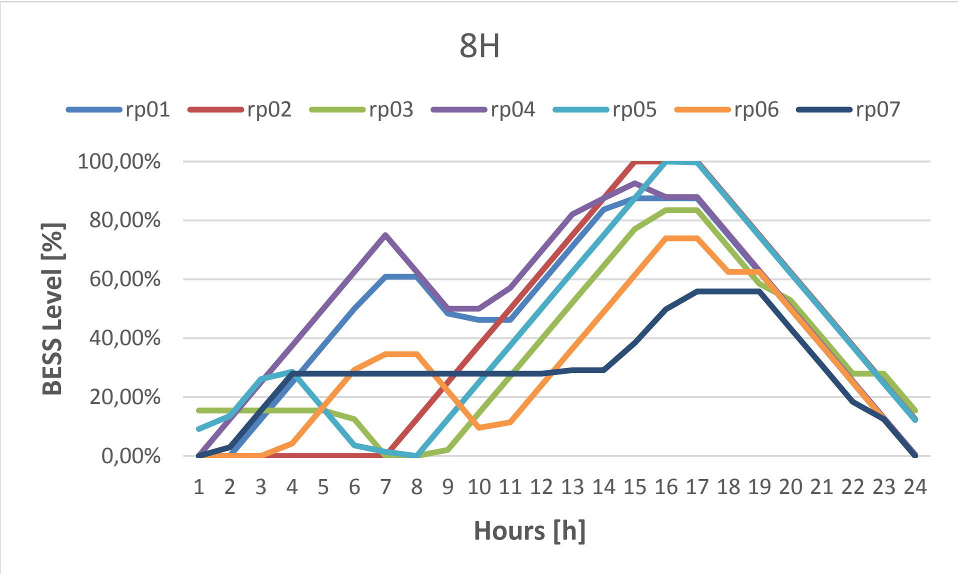


Figure 25.7LRP kmeans BESS level for work cycles of 8h and 16h.

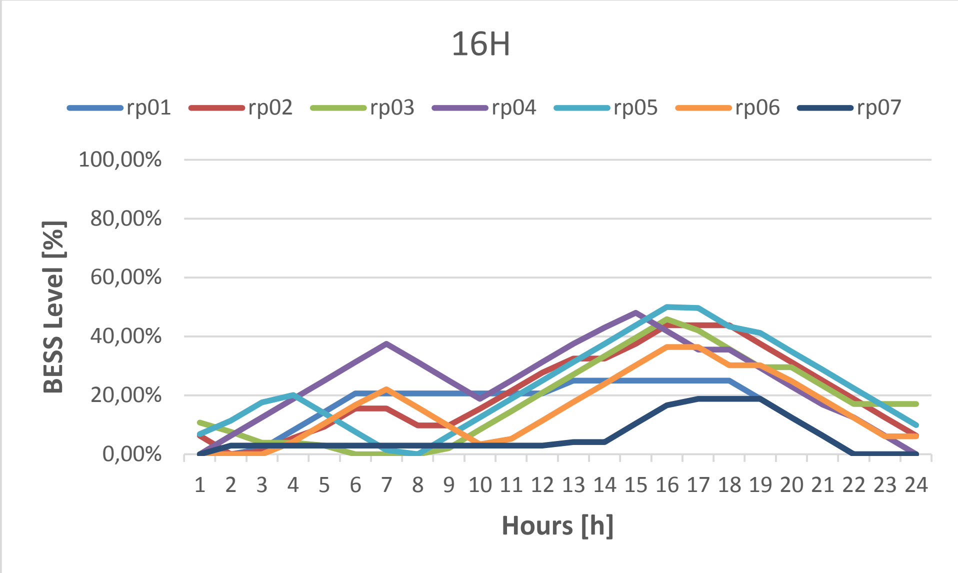


Figure 26. 7LRP kmeans BESS level for work cycle of 32h.

Figure 25 includes a large saturation of the battery capacity in central hours of the day, that can be optimized up until a work cycle of 16h (it has less saturated central hours), included in the same figure. Taking a look at this figure, we estimate the need of having a larger BESS capacity between daylight hours since the batteries mainly charge thanks to the presence of Solar generation. This forces the model not only to operate the system but to include BESS expanding capacity investments. Although this case is pertinent and quite interesting, it is also beyond the scope of this work.

However, Figure 26 displays a 32-hour work cycle that cannot charge BESS levels to its 100%. This is because the BESS total capacity is maintained equal, which results in generating a smaller amount of hourly power as the charge cycle enlarges.

Here, we also try to quantify the impact on the objective function when using longer cycles or the sensibility towards one duration step and another. As it can be seen in Figure 27, all models are economically more feasible as shorter their cycles are: for charges within few hours (2 or 4 hours charging and discharging), the objective function value is lower than for longer charge cycles. The reason behind this result is that faster batteries are able of producing a higher amount of power each hour, fact that slower batteries cannot maintain. The inability of producing great amounts of power impacts on the optimization analysis, resulting in more expensive system total costs.

We must keep in mind that the Full Year Model (Base Case) represents always the 'reality' since it does not consider input data reduction within a year. This is why Chrono model is the most accurate model to shape BESS evolution, especially when considering larger cycles (from 8 hours-that include both the charge and discharge-onwards).

Both 7LRP kmeans and kmedoids models shape this technology worse since they need to specify a representative period that must resemble the battery cycle duration in order for the model to be accurate.

As seen in both Figure 25 and Figure 26, work cycles of 8 and 16 hours are included within the 24 representative hours of 7LRP model so they are well optimized, whether a work cycle of 32 hours are greater than the day registered by 7LRP. In other words, if the 7LRP model specifies a shorter representative period than the total duration of both batteries charge and discharge, it loses hours in which the BESS technology has already begun charging or has not finished its discharge completely. This situation results in a difference between the model and the Base Case, that can be understood as an accuracy error.

It should also be said that the 4-hour BESS charge cycle represents the Base Case presented in Section 3 and this is why 7LRP model behaves in the same way for both clusterization algorithms.

Summing up this results, Figure 27 classifies Chrono as the most accurate model to determine BESS behavior, followed by 7LRP kmeans and 7LRPkmedoids models, if the representative period chosen is, as in this case, a representative week.

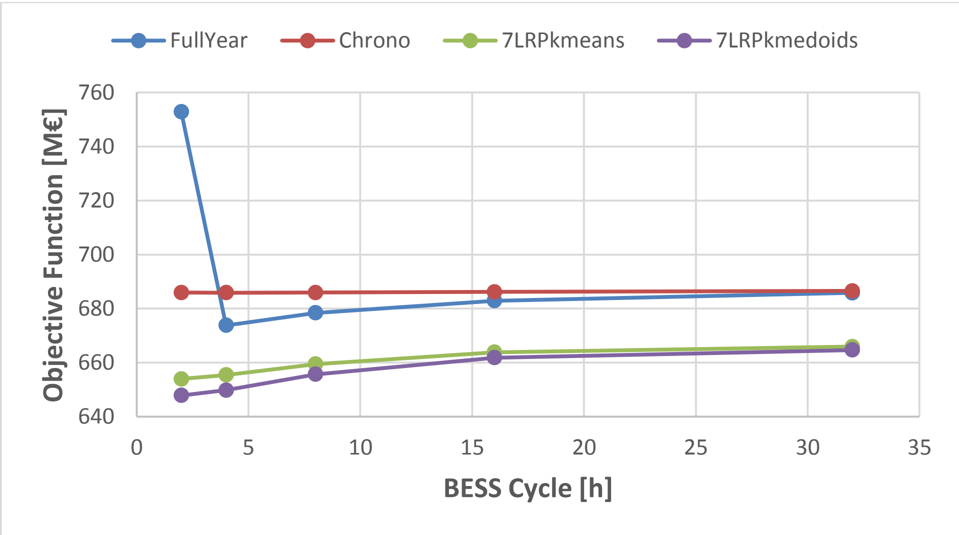


Figure 27. Objective function value – BESS charge cycle duration

Deeply examining Figure 27, we can see that there is a point for all models (16 hours of work cycle: 8 hours for each charge and discharge) where the objective function rises, and we could ask ourselves if BESS technology could be better modelled as a generation technology with a higher inertia, the same Hydro units have.

As a general conclusion, it can be extracted that no model is better than its neighbor, because its functioning mainly depends on the duration of the work cycle and this one depends its either modelled as a fast battery (for work cycles lower than 16 hours) or units working in a larger time grid.

But not only the objective function value is enough to compare the model's behaviors properly. We need to look further at the hourly variation in order to see similarities and differences with full detailed models. This other way of looking at the BESS estimation difference/error is by creating a box and whiskers plot for every model and BESS charge duration, which in fact is resumed in Figure 28.

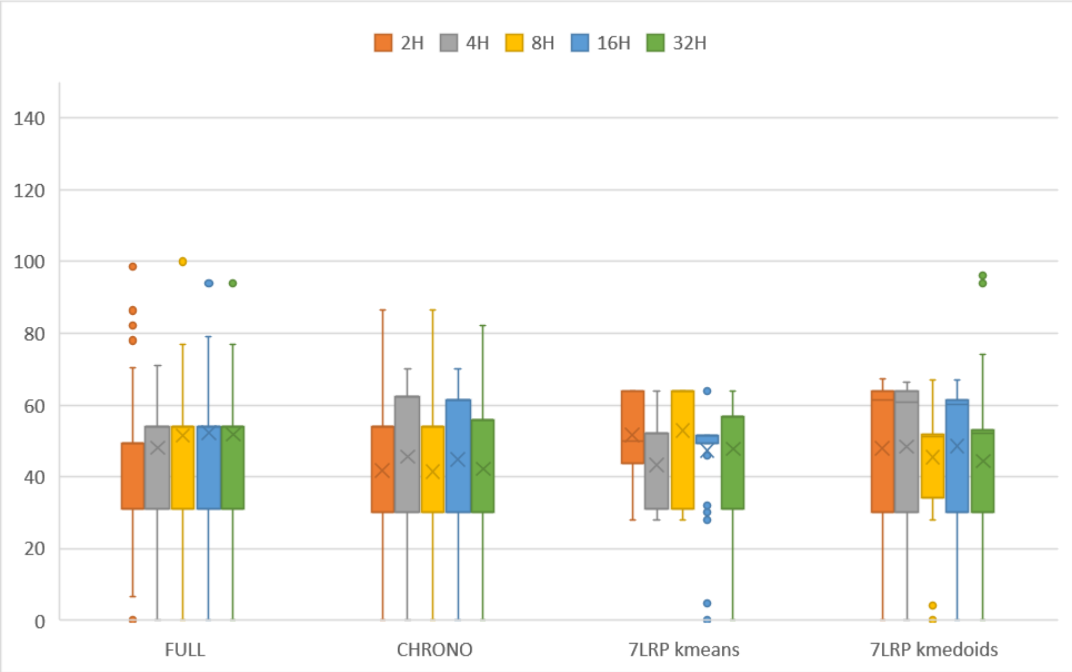


Figure 28. Hourly cost in BESS Analysis[€/MWh]

Figure 28 throws interesting conclusions towards total optimization operation costs: the faster the batteries are, the better for the total system cost since BESS technology will be able of throwing up a great amount of hourly power within a short period of time (it could be completely used up until 6 times with 2 hours charges and up until 3 times in a day with 4 hours ramp up periods).

As it was previously mentioned, the Chrono model lacks from this hourly price deviation which is the same as saying that the objective function remains static and independent from BESS cycle duration. This does not happen for 7LRP model, that has a greater price variability (generally smaller prices) for faster periods than for work cycles greater than the representative period included in 7LRP model, in this case a representative week.

Another important variable to study is the hourly price along the time horizon. As in other variables studied before, the error is reduced as quicker BESS cycles are considered and are generally smaller for the Chrono model if a representative week is chosen in the last 7LRP approximation, fact seen in Figure 29.

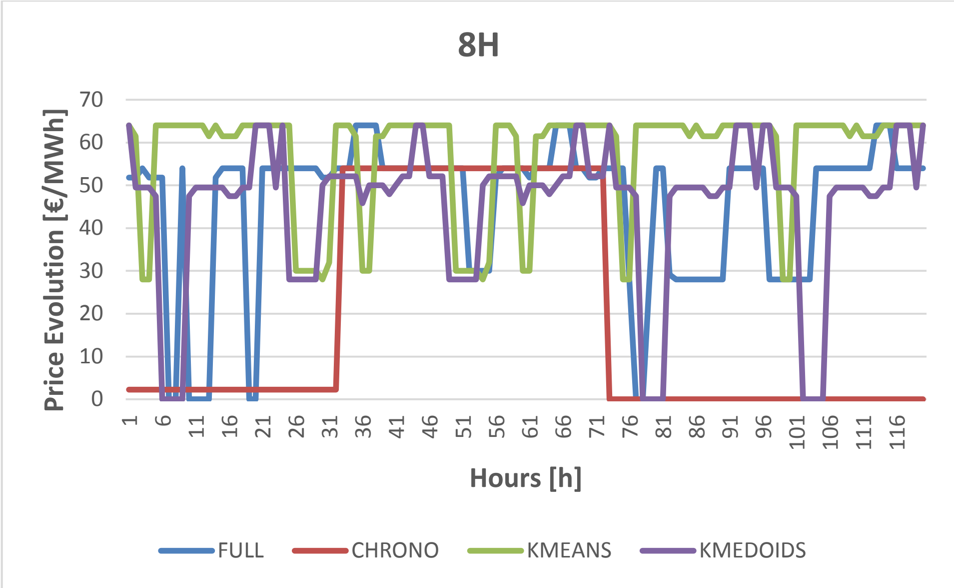
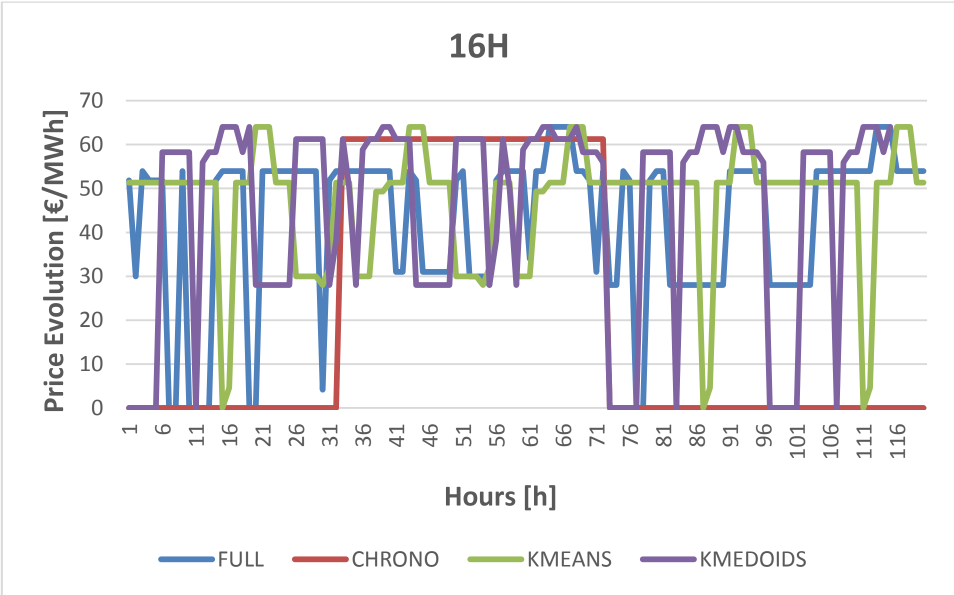


Figure 29. Hourly BESS Price Evolution for all Models in the selected period.



In addition to the objective function value other variables must be studied to come up with a precise approach to the whole, so BESS levels are compared for two different work cycle durations in Figure 30 and Figure 31 for all models. It is important to underline that in order to maintain an accurate comparison between models, the 24 hours included in the representative day chosen by kmeans, have also been picked for the Base Case and the other models. This 24-hour period and adjacent days are pictured in the following figures.

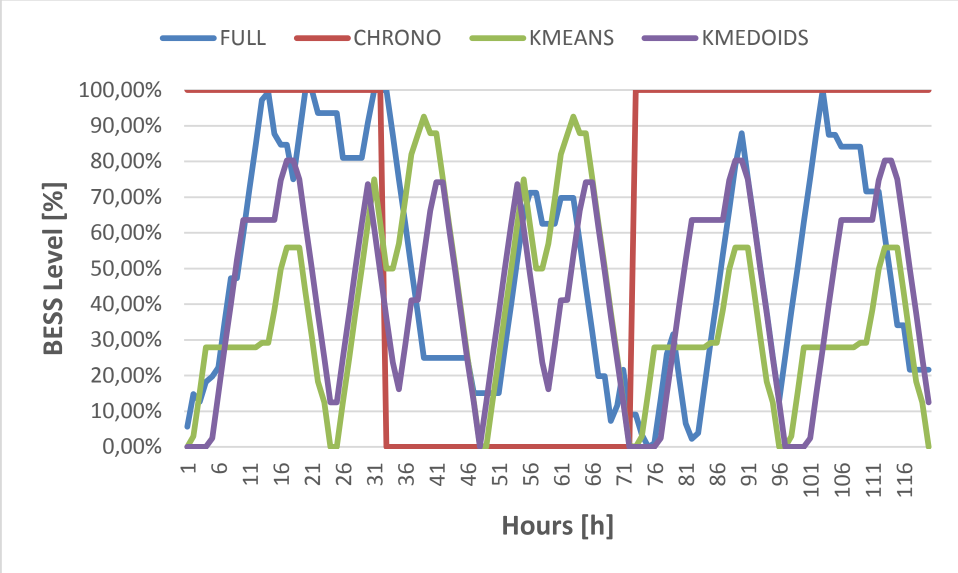


Figure 30. BESS level for a 16h-work cycle.

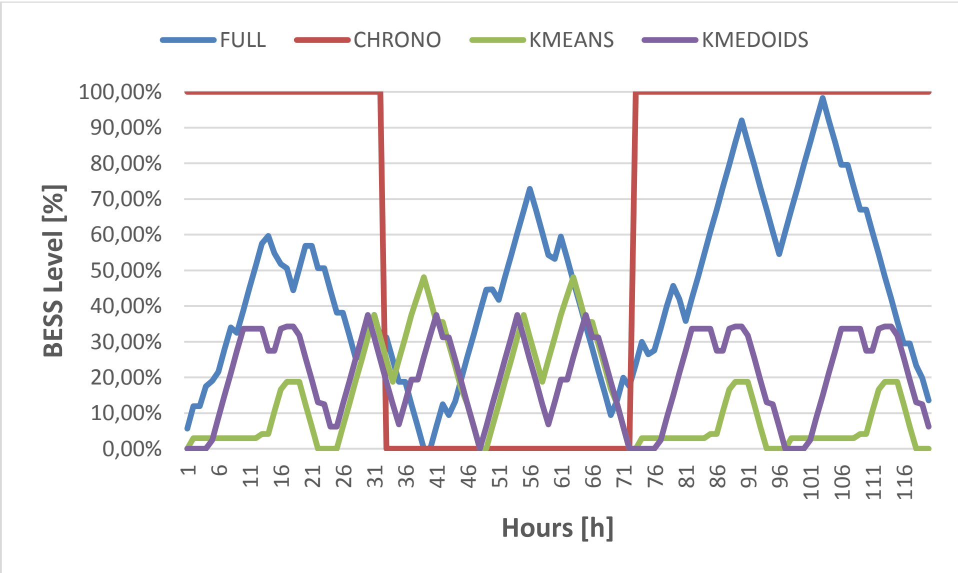


Figure 31. BESS level for a 32h-work cycle.

The battery level evolution throughout these consecutive days reaches the 100% of capacity in a smaller cycle whether the full capacity cannot be reached for those 7LRP models with a representative day as unit, for the reasons previously commented.

Some interesting conclusions of this section are resumed in the following points:

- The batteries' cycles of 8 and 16 hours considered previously saturate the battery capacity in central hours of the day since BESS serve for the storage of solar energy, mainly. If these charge cycles are considered, investment decisions will translate into a big advantage to the system's optimal costs.
- Work cycles above 16 hours are non-optimal since batteries are not able of charging to its maximum capacity due to smaller amounts of energy that could be discharged at each time slice (hours in our optimization problem) as BESS cycles enlarge.

In relation to the impact on the objective function:

- **All models are economically more feasible as shorter their cycles are**: for charges within few hours (4 hours charging and discharging), the objective function value is lower than for longer charge cycles.
- **Chrono model is the most accurate model to shape BESS evolution**, especially when studying larger cycles (from 16h work cycles onwards).
- **7LRP kmeans and kmedoids model BESS worse** because they **need to specify a representative period** that must resemble the battery cycle duration to assure accuracy. If 7LRP's time unit model specifies a representative period similar to the BESS' duration work cycles, the model will behave correctly; if not, an accuracy error will be made.

Also, some remarkable points are included hereunder:

- There is a 'time' point for all models (16 hours of work cycle) where the objective function rises, in which we ask ourselves if BESS technology could be better modelled as a generation technology with a higher inertia, the same Hydro units have.
- As it was previously mentioned, the Chrono model lacks from an hourly price deviation: the objective function remains static and independent from BESS cycle duration. 7LRP model, that has a greater price variability for faster periods than the representative period indicated.
- Regarding the hourly price along the time horizon, the Chrono model is the best placed since it englobes the total price variability of the Full Year model.

References

[1] Tejada-Arango, DA, M. Domeshek and S. Wogrin, "Enhanced Representative Days and System States for Energy Storage Investment Analysis," IEEE, Madrid, Spain, 2018.

[2] S. Pineda and J. Morales, "Chronological Time-Period Clustering for Optimal Capacity Expansion Planning With Storage," IEEE, Málaga, 2018.

[3] H. Teichgraeber and B. AR, "Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison," Elsevier, California, 2018.

[4] M. Boiteux, "La tarification des demandes en pointe: application de la théorie de la vente au coût marginal," Revue Generale d'Electricité, 1949.

Appendix 1 - Notation of Models

Indexes and sets

$t \in \mathcal{T}$	Hourly periods
rp	Representative periods
k	Periods inside a representative period
$h_{t,rp,k}$	Relation among periods and rp
$j \in \mathcal{J}$	Generation technology
$g \subseteq \mathcal{J}$	Thermal generator unit
$s \subseteq \mathcal{J}$	Storage unit
$v \subseteq \mathcal{J}$	Renewable unit
i	Node i
g_i	Generator g connected to node i
$l \in \mathcal{L}$	Lines connecting node i to ii
$li \subseteq \mathcal{L}$	Input lines
$lo \subseteq \mathcal{L}$	Output lines
a_i	actives nodes

Parameters

C^{ENS}	Energy non-served cost [M€/GWh]
C_j^{var}	Slope variable cost [M€/GWh]
C_j^{int}	Intercept variable cost [M€/h]
C_j^{st}	Start-up cost [M€]
$D_{rp,k,i}$	Hourly demand per node [GW]
$D_{rp,k,s}$	Inflows for hydro storage [GWh]
ε_g	EFOR: interruptibility of generation unit
E_s	Pumping Efficiency [p.u.]
P_j^+	Maximum output production [GW]
P_j^-	Minimum output production [GW]
RU_j	Ramp up limit [GW]
RD_j	Ramp down limit [GW]
W_{rp}^{rp}	Representative periods weight [h]
W_{rp}^{k}	Hourly weight for each rp [h]
M	Moving window for inter-period [h]
N	Transmission network [Y/N]
$R0_s$	Initial reserve [MWh]
R_s^+	Maximum Reserve [MWh]
R_s^-	Minimum Reserve [MWh]
Q_s^+	Maximum Consumption [MW]

Variables

$u_{rp,k,g}$	Binary commitment decision
$y_{rp,k,g}$	Binary start-up decision
$z_{rp,k,g}$	Binary shutdown decision

$p_{rp,k,j}$	Production of the unit [GW]
$q_{rp,k,s}$	Consumption of the unit [GW]
$\hat{p}_{rp,k,g}$	Production above min [GW]
$pns_{rp,k,i}$	Power non-served [GW]
$pf_{rp,k,l}$	Power flow
$sp_{rp,k,s}$	Spillage [GWh]
$R_{rp,k,j}^{inter}$	Reserve at the end inter period [GWh]
$R_{rp,k,j}^{intra}$	Reserve at the end intra period [GWh]
$R0_s$	Initial reserve

Appendix 2 – Generation mix

Firstly, we can find the details of the fossil fuels generation plants in the system. Nuclear, thermal, combined cycle, fueloil and a gas turbine are the main part of the technologies used. Furthermore, the specification of the hydro reservoir and the batteries are defined.

NUCLEAR

Amount of plants: 1

Nuclear:		1
P_j^+	Maximum output production [MW]	771.6
P_j^-	Minimum output production [MW]	771.6
RU_j	Ramp up limit [MW]	0
RD_j	Ramp down limit [MW]	0
C^{ENS}	Energy non-served cost [M€/GWh]	10000
C_j^{var}	Slope variable cost [M€/GWh]	0.015
C_j^{int}	Intercept variable cost [M€/h]	0
C_j^{st}	Start-up cost [M€]	0
ε_g	EFOR: interruptibility of generation unit	0

COAL

Amount of plants: 4
(Coal, Brown Lignite, Imported Coal SubBituminous, Imported Coal Bituminous)

Coal:		1	2	3	4
P_j^+	Maximum output production [MW]	588.0	203.1	150.4	194.4
P_j^-	Minimum output production [MW]	235.2	81.2	60.2	77.8
RU_j	Ramp up limit [MW]	88.2	30.5	22.6	29.2
RD_j	Ramp down limit [MW]	88.2	30.5	22.6	29.2
C^{ENS}	Energy non-served cost [M€/GWh]	10000	10000	10000	10000
C_j^{var}	Slope variable cost [M€/GWh]	0.054	0.052	0.052	0.05
C_j^{int}	Intercept variable cost [M€/h]	0.001	0.001	0.001	0.001
C_j^{st}	Start-up cost [M€]	0.04	0.04	0.04	0.04
ε_g	EFOR: interruptibility of generation unit	0	0	0	0

CCGT

Amount of plants: 4

CCGT		1	2	3	4
P_j^+	Maximum output production [MW]	500.0	500.0	500.0	667.5
P_j^-	Minimum output production [MW]	100.0	100.0	100.0	133.5
RU_j	Ramp up limit [MW]	200.0	200.0	200.0	267.0
RD_j	Ramp down limit [MW]	200.0	200.0	200.0	267.0

C^{ENS}	Energy non-served cost [M€/GWh]	10000	10000	10000	10000
C_j^{var}	Slope variable cost [M€/GWh]	0.03	0.031	0.034	0.028
C_j^{int}	Intercept variable cost [M€/h]	0.009	0.009	0.009	0.009
C_j^{st}	Start-up cost [M€]	0.03	0.03	0.03	0.03
ε_g	EFOR: interruptibility of generation unit	0	0	0	0

OCGT

Amount of plants: 3

OCGT:		1	2	3
P_j^+	Maximum output production [MW]	400.0	400.0	400.0
P_j^-	Minimum output production [MW]	0	0	0
RU_j	Ramp up limit [MW]	400.0	400.0	400.0
RD_j	Ramp down limit [MW]	400.0	400.0	400.0
C^{ENS}	Energy non-served cost [M€/GWh]	10000	10000	10000
C_j^{var}	Slope variable cost [M€/GWh]	0.064	0.067	0.07
C_j^{int}	Intercept variable cost [M€/h]	0.003	0.003	0.003
C_j^{st}	Start-up cost [M€]	0	0	0
ε_g	EFOR: interruptibility of generation unit	0	0	0

FUELOILGAS

Amount of plants: 1

P_j^+	Maximum output production [MW]	441.8
P_j^-	Minimum output production [MW]	0
RU_j	Ramp up limit [MW]	441.8
RD_j	Ramp down limit [MW]	441.8
C^{ENS}	Energy non-served cost [M€/GWh]	10000
C_j^{var}	Slope variable cost [M€/GWh]	0.123
C_j^{int}	Intercept variable cost [M€/h]	0.018
C_j^{st}	Start-up cost [M€]	0.06
ε_g	EFOR: interruptibility of generation unit	0

STORAGE: HYDRO

Amount of plants: 1

P_s^+	Maximum output production [MW]	600.0
P_s^-	Minimum output production [MW]	0
$R0_s$	Initial reserve [MWh]	750000
R_s^+	Maximum Reserve [MWh]	960000
R_s^-	Minimum Reserve [MWh]	300000

STORAGE: BATTERIES

Initial amount of plants: 1

P_s^+	Maximum output production [MW]	200.0
P_s^-	Minimum output production [MW]	0
E_s	Pumping Efficiency [p.u.]	0.96
R_s^+	Maximum Reserve [MWh]	800
Q_s^+	Maximum Consumption [MW]	200

RES: WIND

Amount of plants: 1

P_r^+	Maximum output production [MW]	2900.0

RES: SOLAR

Amount of plants: 1

P_r^+	Maximum output production [MW]	2000.0

Appendix 3 – rMIP Tables and Figures

rMIP General RES Sensitivity Analysis

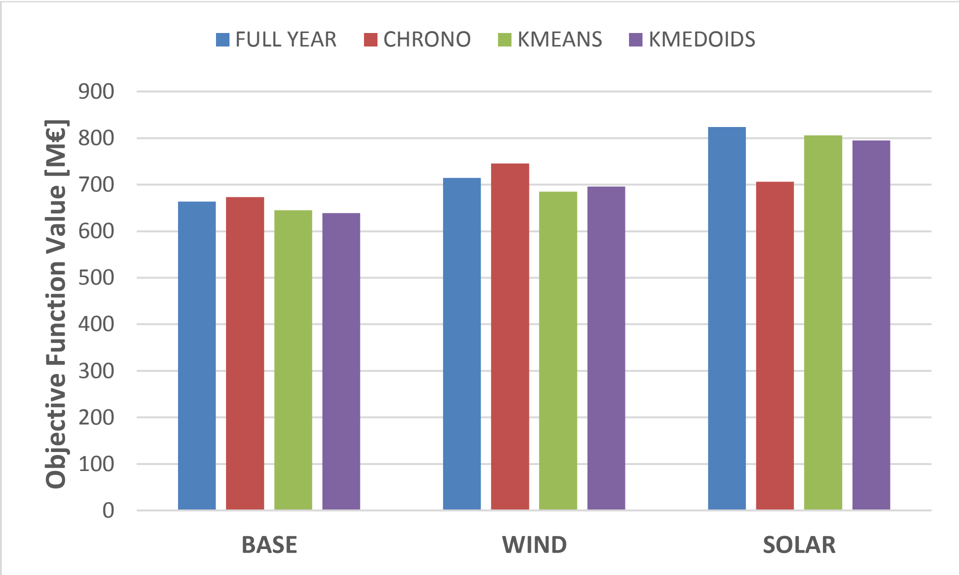


Figure 32. Objective Function for the three RES cases.

	Chrono	7LRP kmeans	7LRP kmedoids
Base Case	1.46%	-2.84%	-3.70%
All Wind	4.34%	-4.09%	-2.60%
All Solar	-14.28%	-2.21%	-3.50%
RMSE	**15.00%**	**5.45%**	**5.72%**

Table 17. Model and Case objective function Error

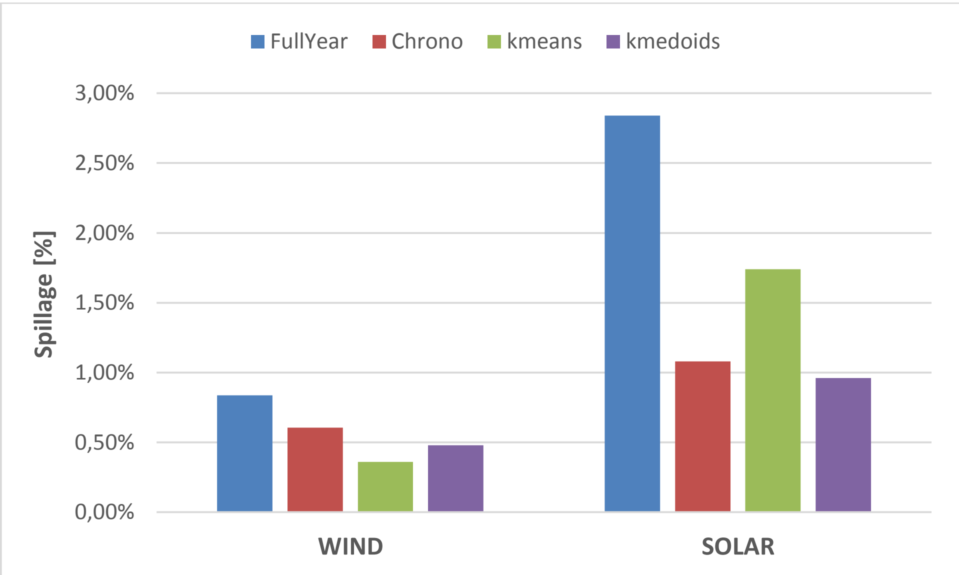


Figure 33. RES Spillage.

rMIP WIND Sensitivity Analysis

	FULL YEAR	CHRONO	7LRP kmeans	7LRP kmedoids
Objective function [M€]	714,365	745,388	685,152	695,825
Objective Function Value Difference [%]	**7,63%**	**10,68%**	**6,24%**	**8,87%**
Number of discrete variables	340704	6552	6552	6552
Number of equations	672673	12937	12873	12873

Table 18. Summary of results: RES Wind production.

	CHRONO	**7LRP KMEANS**	**7LRP KMEDOIDS**
Nuclear	-5,34%	5,44%	4,63%
Fuel Oil Gas	0,00%	0,00%	0,00%
BESS	-81,73%	-12,19%	3,77%
Wind	-0,29%	1,24%	0,72%
Solar	0,00%	0,00%	0,00%
Coal	-66,21%	-100,00%	-58,21%
CCGT	3,70%	0,74%	-0,67%
OCGT	713,12%	100,00%	-79,67%
Hydro	0,00%	0%	-4,15%
Total	-0,96%	-0,14%	-0,33%
RMSE	720,87%	142,06%	98,94%

Table 19.Inter Model Comparison: RES Wind Production.

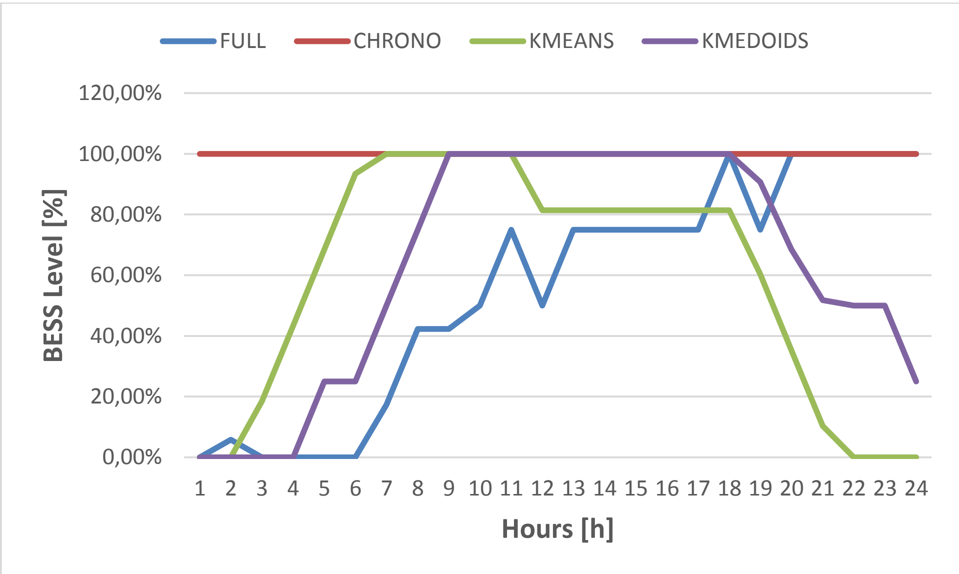


Figure 34. Inter Model Comparison: BESS Level

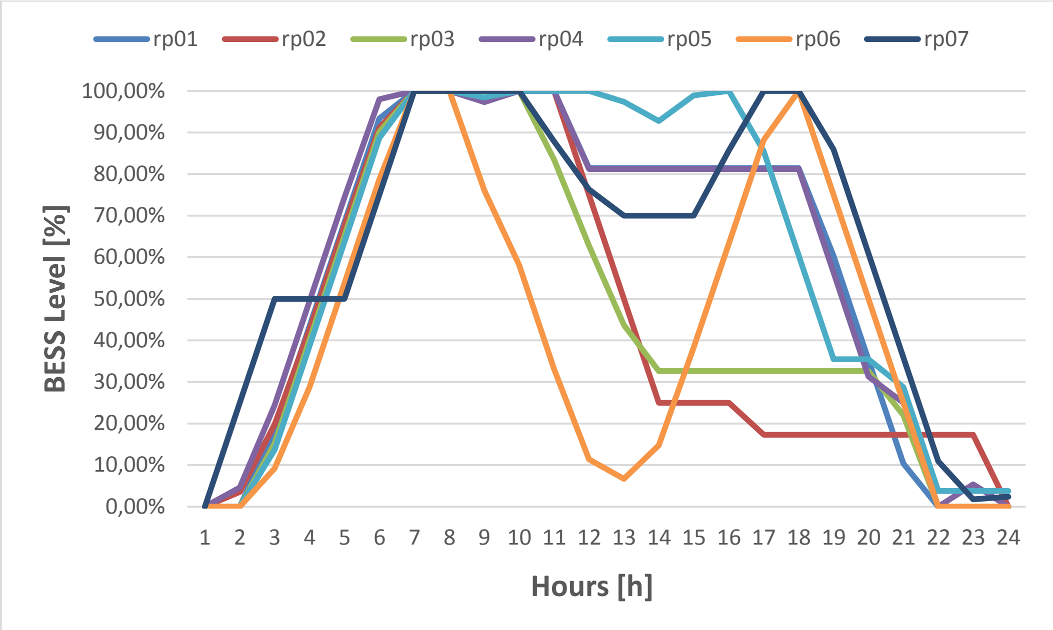


Figure 35. BESS Evolution: 7LRP kmeans RES Wind production.

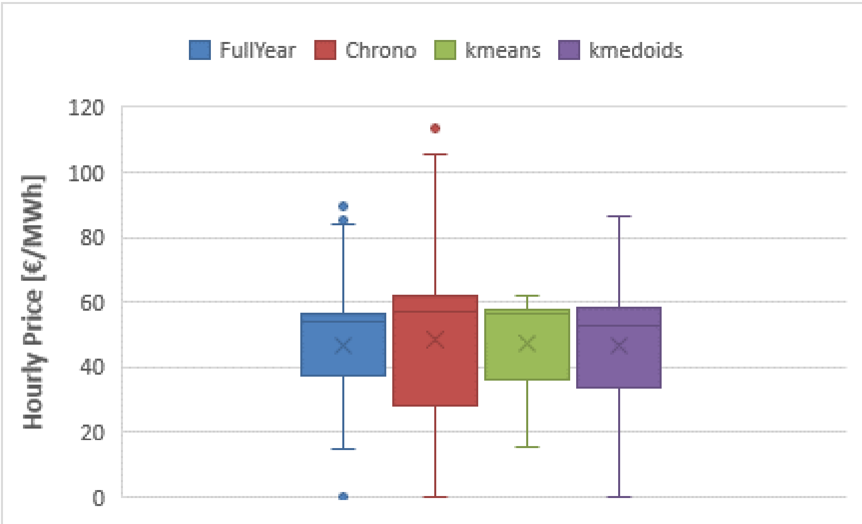


Figure 36. Hourly price evolution: RES Wind Production.

rMIP SOLAR RES Sensitivity Analysis

	FULL YEAR	CHRONO	7LRP kmeans	7LRP kmedoids
Objective function [M€]	824,101	706,387	805,899	795,284
Objective Function Difference [%]	**24,16%**	**4,89%**	**24,96%**	**24,43%**
Number of discrete variables	340704	6552	6552	6552
Number of equations	672673	12937	12873	12873

Table 20. Summary of results: RES Solar production

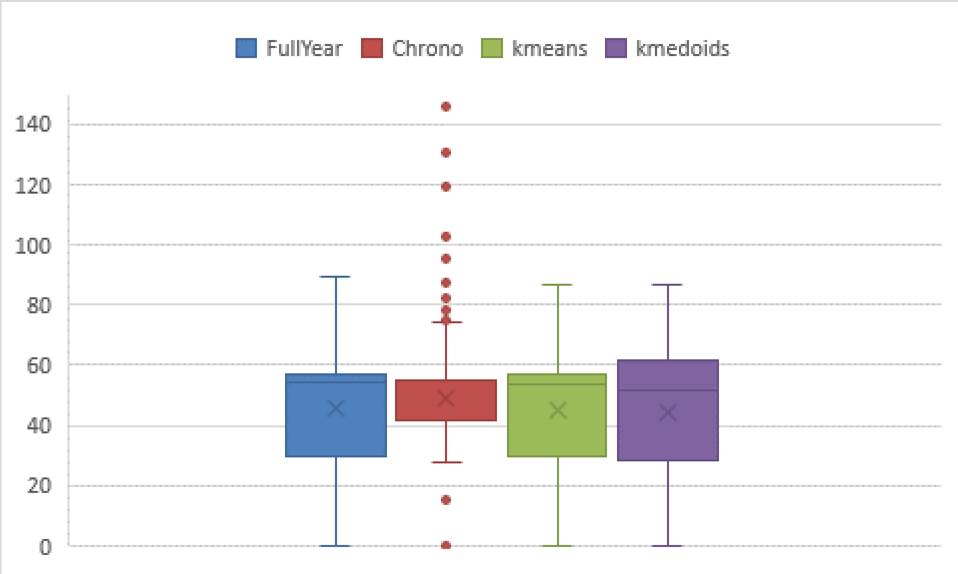


Figure 37. Hourly price evolution: RES Solar Production

	CHRONO	**7LRP KMEANS**	**7LRP KMEDOIDS**
Nuclear	15,45%	0,13%	-1,65%
Fuel Oil Gas	0,00%	0,00%	0,00%
BESS	-87,50%	-12,32%	12,42%
Wind	0,00%	0,00%	0,00%
Solar	4,32%	3,24%	5,61%
Coal	-92,99%	-31,88%	-24,75%
CCGT	-1,45%	1,09%	-1,20%
OCGT	88,64%	-35,62%	-40,21%
Hydro	0,00%	0,00%	-3,78%
Total	-1,09%	-0,15%	-0,21%
RMSE	156,27%	49,49%	49,33%

Table 21. Inter-Model Comparison: RES Solar Production.

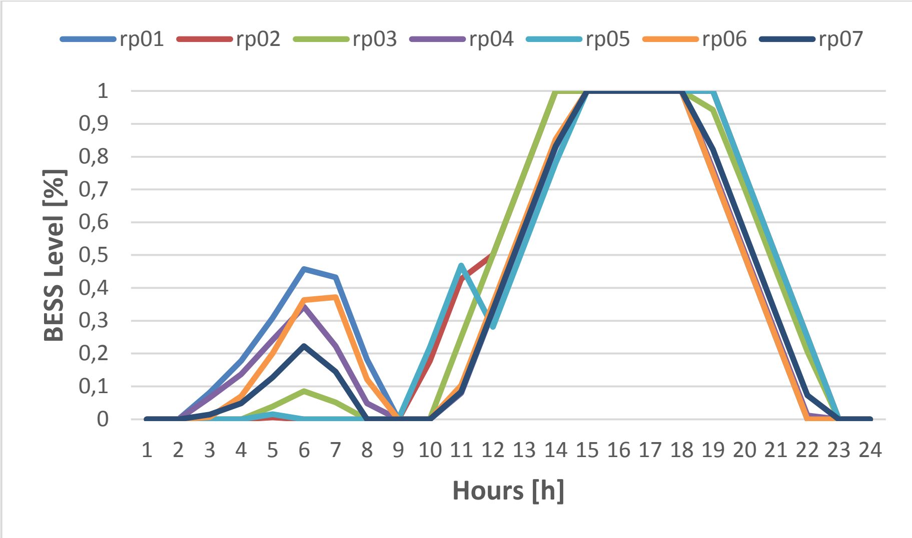


Figure 38. BESS Evolution: 7LRP kmeans RES Solar production.

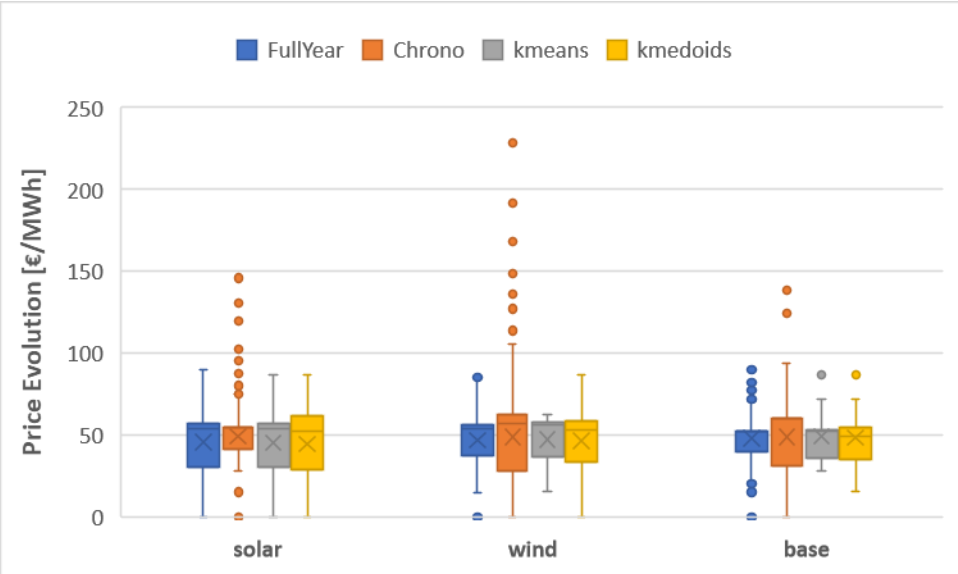


Figure 39. rMIP Transversal RES Analysis.

rMIP BESS Sensitivity Analysis

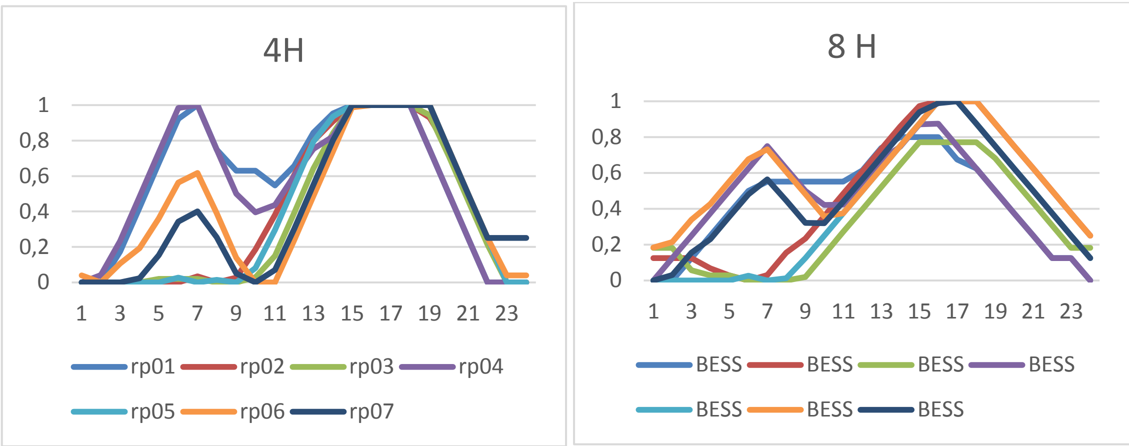
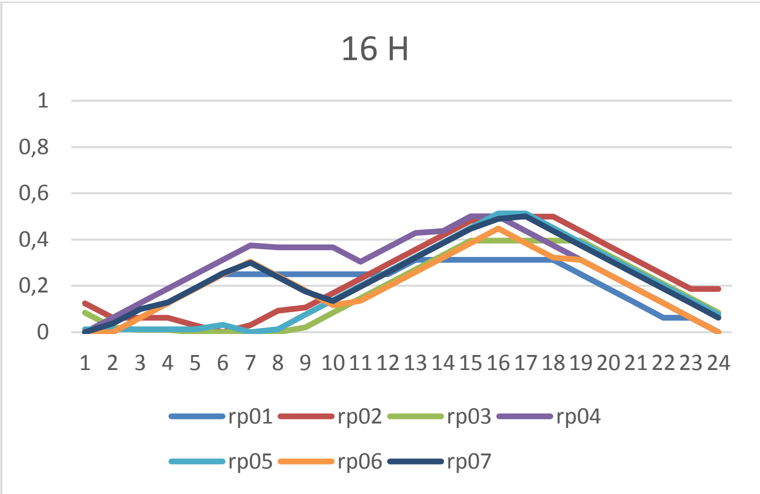


Figure 40.7LRP kmeans BESS level for work cycles of 8h and 16h.



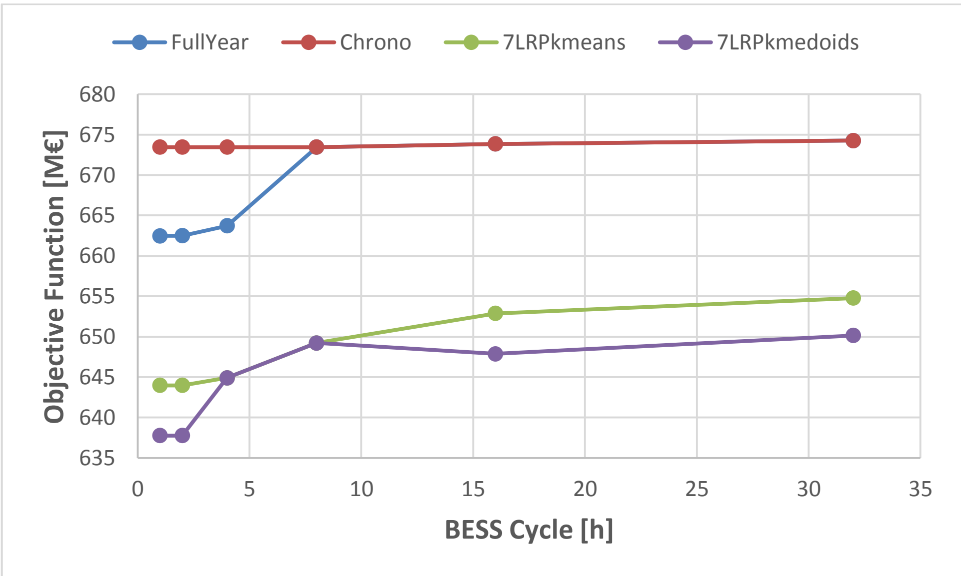


Figure 41. Objective function value – BESS charge cycle duration

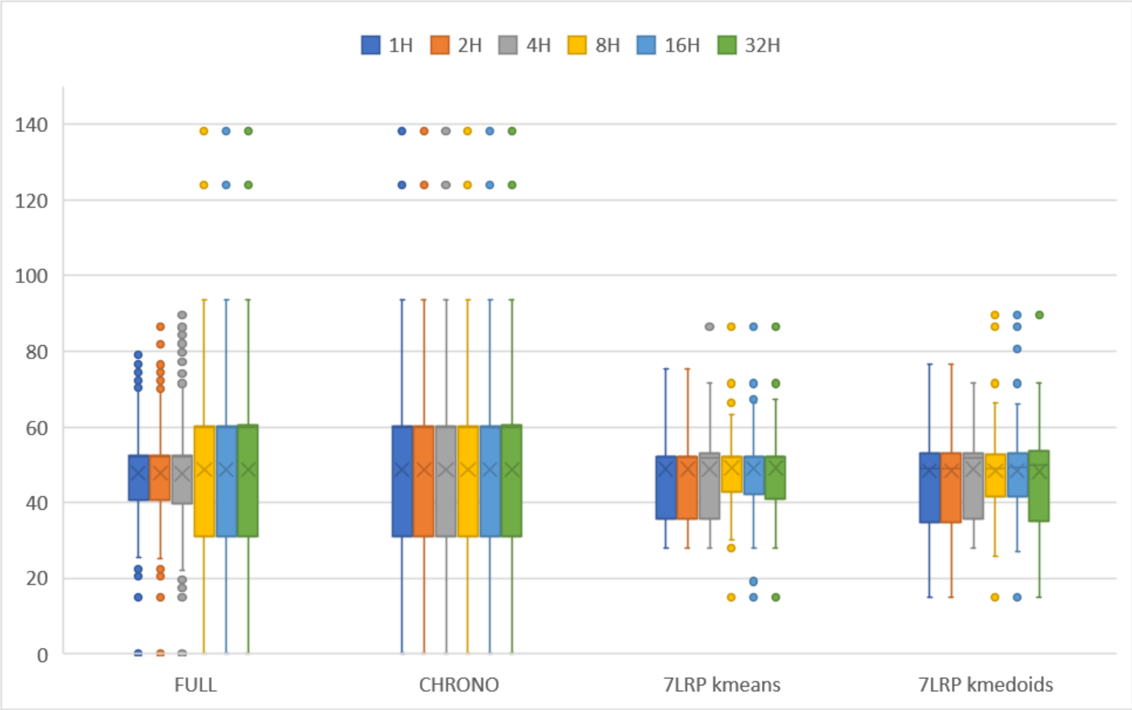


Figure 42. Hourly cost in BESS Analysis[€/MWh]

Appendix 4 – Tools used

Matlab R2018b

This software has been used for all operations that handled inputs. The most important function to describe is 'kmeans' and 'kmedoids' which are in charge of iterating the information and create the clusters:

```
TDur = readtable('InputMatrix.xlsx','sheet','Table2');
NumClusters = 7;
a = zeros(Months);
Days = zeros(Months);
  for k = 1:Months
      a(k) = sum(TDur.Duration(1:k))/h;
      Days(k) =    TDur.Duration(k) /h ;
      if k==1
        b = 1;
       else
        b = a(k-1) + 1 ;
      end

     [Hindex, RD, sumd, Dist] = ...

               kmeans(DaysData(b:a(k),:),  NumClusters,'Replicates',
...

           1000,'MaxIter',10000,'Display','final');


     opts = statset('Display','iter');

     [Hindex,RD, sumd, Dist, midx, info] = ...

           kmedoids(DaysData(b:a(k),:), NumClusters, ...

           'Replicates',1000 ,'Options',opts);
```

GAMS: CPLEX 12.6

The commercial software used to optimize our model used CPLEX 12.6 as solver. All solutions have been achieved through MIP modelling.

This solver allows to assemble simple models and complex ones and arrive to a solution through the following characteristics:

- Branch and cut algortihms solving LP subproblems first.
- Sets of semi-continuous and semi-integer variables.
- RMIP problems can also be solved to bear out the optimized solution.

For new formulations a very useful tool for CPLEX has been establishing the epgap. To check if the model is correctly defined the first solution reached is little away from the optimized one (0.5% from the optimal). This approximated solution is reached faster and allows to correct errors first before executing the model to its optimal.