This article is an accepted version. Please cite the published version: DOI https://doi.org/10.1016/j.renene.2025.122730

Efficient Hydropower Modeling for Medium-Term Hydrothermal Planning Using Data-Driven Approaches

Jesús D. Gómez-Pérez[a,*], Francisco Labora[a], Jesus M. Latorre-Canteli[b] and Andres Ramos[a]

[a]*Comillas Pontifical University, Institute for Research in Technology, Calle Alberto Aguilera, 23, Madrid, 28015, Spain*
[b]*Iberdrola Generación, Tomás Redondo, 1, Madrid, 28033, Spain*

ARTICLE INFO

Keywords:
Equivalent hydropower plants
Mid-term planning models
K-means
Linear Regression models
Fourier series filtering
Ridge regularization
Linear optimization models

ABSTRACT

The continuous rise of renewable energy in the global energy mix highlights the need to analyze and enhance traditional energy plants' flexibility to support integration. Hydropower, with its rapid response capabilities and significant energy storage, plays a vital role in this context. However, simplifications are required due to the complex interconnections among cascaded hydropower plants and the inherent uncertainty of water inflows. This study presents a data-driven methodology for representing hydropower plants physically and through equivalent energy models, accounting for inflow uncertainties implicitly. Using historical data, we apply analytical techniques—including auxiliary linear models, load-duration curves, and filtering methods in linear regressions—to configure key hydropower parameters such as water inflows, reservoir boundaries, and hydropower plant production limits. These methods can be applied across hydro systems of different scales. We have validated our approach for the Spanish system for 2019 and 2025, demonstrating its efficacy.

Nomenclature

Sets (calligraphic)

$\mathcal{P}$ Set of pumping plants. $p \in \mathcal{P}$

$\mathcal{R}$ Set of reservoirs. $r \in \mathcal{R}$

$\mathcal{R}_r$ Set of tuples $(\underline{r}, r)$ that relate the upstream reservoir $\underline{r}$ and reservoir r

$\mathcal{R}_p$ Set of tuples (r, p) that relate the pumping plant p and the reservoir r where the water is pumped

$\mathcal{P}_r$ Set of tuples (p, r) that relate the pumping plant p and reservoir r where the water is taken from

Parameters (uppercase)

RV_r Volume level of reservoir r at the end of the day. $[hm^3]$

RVI_r Volume level of reservoir r at the start of the day. $[hm^3]$

OUT_r Water outflows of reservoir r throughout the day. $[hm^3]$

PU Basin pumping consumption. $[MWh]$

EC_p Production function of pumping plant p. $[kWh/m^3]$

EFF_p Round-trip efficiency of pumping plant h

PN, $PN2$ Penalty values.

$\underline{PM}_p$, $\overline{PM}_p$ Lo. and Up. bound of each pumping p. $[MW]$

Variables (lowercase)

if_r Natural water inflows in reservoir r throughout the day. $[hm^3]$

aif_r Slack variable to avoid infeasibilities caused by violations in the lower bound of reservoir r. $[hm^3]$

of_r, $of_{to(r)}$ Water outflows resulting from electricity production in (into a) reservoir r throughout the day. $[hm^3]$

pm_p Consumption of a pumping plant p. $[MW]$

pmw_r Pumped water from a reservoir r. $[hm3]$

$pmw_{to(r)}$ Pumped water to a reservoir r. $[hm3]$

sp_r, $sp_{to(r)}$ Spillage from (to) a reservoir r throughout the day. $[hm^3]$

ir_r Irrigation from a reservoir r throughout the day. $[hm^3]$

1. Introduction

Hydropower technology is essential for enhancing power system flexibility due to its rapid response and significant energy storage capabilities. The growing integration of renewable energy sources underscores the importance of hydropower, particularly in addressing the inherent variability of these resources Zhang, Li, Chen, Xu and Mahmud (2021).

Comprehensive modeling of hydropower is crucial for short-, medium-, and long-term energy planning. This modeling must encompass system physics, dynamics, uncertainty, and other water uses. From a physics point of view, key parameters typically included are minimum and maximum power outputs, ramping capabilities, storage limits, cascaded hydropower plant topologies, and functions that model the conversion of water potential energy into electrical energy [Niu and Insley (2013), Stoll, Andrade, Cohen, Brinkman and Brancucci Martinez-Anido (2017)]. Regarding the system's dynamics, it is crucial to represent the variability of renewable resources and daily operational cycles by selecting an appropriate time step (hourly, daily, weekly) Hoffmann, Priesmann, Nolting, Praktiknjo, Kotzur and Stolten (2021). Additionally, uncertainty modeling, involving stochastic

This work results from the research project funded by the AVANHID project. Grant CPP2021-009114 funded by MCIN/AEI/ 10.13039/501100011033 and by "European Union NextGenerationEU/PRTR"

*Corresponding author
jdgomez@comillas.edu (J.D. Gómez-Pérez); francisco.labora@iit.comillas.edu (F. Labora); jesuslc@comillas.edu (J.M. Latorre-Canteli); andres.ramos@comillas.edu (A. Ramos)
ORCID(s): 0000-0002-0510-3442 (J.D. Gómez-Pérez); 0009-0008-5686-0292 (F. Labora); 0000-0002-1227-5453 (J.M. Latorre-Canteli); 0000-0002-8871-1872 (A. Ramos)

models and optimization techniques, is essential to address the uncertainty of water inflows Yue, Pye, DeCarolis, Li, Rogan and Gallachóir (2018). Finally, hydropower operations are also influenced by non-electrical water uses such as water supply, recreation, and industry, which must be considered in the models Stoll et al. (2017).

Different conclusions regarding hydropower have been drawn in the literature. Among the most significant findings are the following: stochastic models are deemed the most appropriate due to the inherent uncertainty of water inflows Muhammad and Pflug (2014). Furthermore, researchers have found that employing weekly or monthly time steps is suitable for representing system dynamics Cerisola, Latorre and Ramos (2012). These models necessitate the inclusion of storage limits, production limits, upward and downward ramps, and production functions for accurate representation Stoll et al. (2017). Recent developments have underscored the need to reduce the time step of models owing to the variability of renewable resources. Hourly time steps have emerged as the most suitable option for these applications Ringkjøb, Haugan and Solbrekke (2018). The need to optimize and simulate complete years has further necessitated various simplifications, such as adopting equivalent hydropower representations and deterministic models.

The concept of hydropower equivalents has been explored in various studies. In [Arvanitidits and Rosing (1970) and Arvanitidits and Rosing (1970)], the authors proposed a composite model for multi-reservoir systems based on the potential energy of water in each reservoir. This model aggregates reservoirs, considering that the water within each reservoir can generate energy through its hydropower plant and any downstream hydropower plants that leverage the flow of this water. However, it did not account for certain features like ramps or pumping stations, leading to high flexibility when integrating into optimization frameworks.

The analysis conducted in de Amezúa (2003) builds upon the aggregation approach proposed in Arvanitidits and Rosing (1970). This study aggregated hydropower plants into programming units (PrU) and included pumping units in calculating natural inflows. It is important to highlight that this work employed a deterministic model, and additional considerations were not incorporated to handle the water inflow uncertainty.

In González, Villar, Díaz and Campos (2013), the authors used linear models to represent Spanish hydropower, aggregating plants into conventional storage and closed-loop pumped storage types. However, this approach did not account for temporal correlations or seasonality, potentially leading to imprecise reserve estimations.

In the review de Queiroz (2016), the impact of equivalent hydropower plants on the performance of stochastic models is emphasized. However, the author highlights that this approach may affect system operational decisions, leading to undesirable errors stemming from approximations in the problem formulation.

In Härtel and Korpås (2017), the authors developed more accurate hydropower equivalents by considering different topologies of cascaded hydropower plants and introducing synthetic reservoirs to represent pumping plants. This work presents a more precise equivalent representation of hydropower systems. However, it did not incorporate additional conditions to represent the uncertainty of water inflow.

In Löschenbrand and Korpås (2017), the authors proposed a multi-objective approach using genetic algorithms to generate aggregated equivalents, maximizing similarity to the original system. It is important to note that this study considers a single scenario; therefore, it lacks consideration for reducing the flexibility of the equivalent hydropower due to the water inflow uncertainty. In Blom, Söder and Risberg (2020), the authors extended this work by including multi-scenario models, resulting in more robust equivalents. Continuing their research, the authors in Blom and Söder (2022) compare different techniques for solving the bilevel problem introduced in their previous work. This comparative analysis aims to identify the most effective approach for generating equivalent hydropower plants. In Blom and Söder (2024), the authors apply the Karush-Kuhn-Tucker (KKT) conditions to the bilevel problem and utilize McCormick envelopes and a modified Benders method to solve it faster than previous methods.

Finally, authors in Helseth and Mo (2022) employed Stochastic Dual Dynamic Programming to represent a hydropower equivalent.

The recent studies mentioned above have improved the methodologies for generating aggregated hydropower equivalents, but they often involve complex applications and a strong reliance on underlying programming models. Our approach diverges by prioritizing data dependency and enhancing customizability while managing the flexibility of the resulting equivalent systems. We apply statistical methods to establish appropriate parameter ranges for hydropower equivalents, allowing for more adaptable and data-driven representations that simplify implementation and improve optimization model robustness.

The main contributions of this paper are:

- Proposal of a data-driven framework for modeling hydropower energy systems, considering factors such as natural inflows, production limits, storage, ramping capabilities, and intra-basin interactions.
- Introduction of analytical methods, including auxiliary linear programs, clustering techniques, and Fourier series decomposition filtering, to derive hydropower parameters considering diverse scenarios.
- Provision of a realistic database of the Spanish energy system for medium-term studies, focusing on hydropower modeling.
- Analysis of the Spanish system's operation for 2025 to illustrate the effectiveness of the proposed approach, highlighting the behavior and performance of the introduced framework.

2. Open Data in Spain

This section highlights the available web-based data in Spain, emphasizing its potential for deriving the hydropower plant parameters needed in optimization models.

Civil engineering data. MITECO

The Spanish Ministry of Ecological Transition and Demographic Challenge (MITECO) provides a comprehensive database on reservoirs and dams in Spain, which helps obtain basin and reservoir parameters.

- *Reservoir and dam database [MITECO RV]:* Contains information on maximum and effective volumes, net-head for power plants, and interconnections between hydropower plants and reservoirs[1]. This database lists 374 reservoirs in mainland Spain, including 93 hydropower reservoirs with capacities ranging from 5 to 3160 hm^3.
- *GIS map [MITECO MAP]:* Geographical locations of all rivers and reservoirs in Spain are publicly available through layer files, essential for establishing topological relationships within a basin[2].
- *CEDEX - Aflíqe [CEDEX]:* The Center for Studies and Experimentation in Civil Engineering (CEDEX) provides daily reports on reservoir volumes and outflows, enabling the inference of total daily inflows. However, determining natural inflows requires a more nuanced methodology, explained in section 3.1.
- *Hydro Bulletin [MITECO]:* MITECO conducts topographical studies to define volume-height curves for reservoirs. Although this information is not publicly available, MITECO does publish a weekly report for each hydro basin that includes reserve levels in water and energy.

Power system data. e-SIOS

The System Operator Information System (e-SIOS) offers detailed data on Physical Units (PhUs) and Programming Units (PrUs), their historical power production, and rated capacities, which are useful for obtaining hydropower plant parameters.

- *List of Physical Units (PhU) and Programming Units (PrU) [ESIOS UF and ESIOS UP]:* Includes information on hydropower plants and their grouping into programming units for market participation. Mainland Spain has 1,282 physical hydropower plants ranging from 0.1 to 239 MW, grouped into 256 programming units ranging from 0.1 to 3,543 MW. Among these, 111 physical hydropower units have capacities exceeding 50 MW, and there are 35 programming units with capacities over 50 MW.
- *Pumping and production [ESIOS]:* ESIOS provides hourly values on power production and pumping consumption per programming unit, which helps enhance the estimation of natural inflows.

[1] Iberdrola (2006) provides comprehensive insights into Iberdrola's hydropower plants.

[2] It is important to note that we use Google Maps and Google Earth to validate some data.

3. Proposed Methods for Hydropower Parameter Extraction

This section explores the analytical methods used for parameter extraction applicable to both physical and equivalent aggregated energy representations of hydropower plants. The physical representation captures the detailed dynamics of the hydropower system, including water flow, reservoir levels, and turbine operations. In contrast, the equivalent aggregated energy model simplifies these dynamics into energy terms, facilitating integration with broader energy system models. These methods play a key role in the methodology outlined in Figure 1, which details the extraction of hydropower parameters, their validation, and the subsequent application in a real-world case study in Spain.

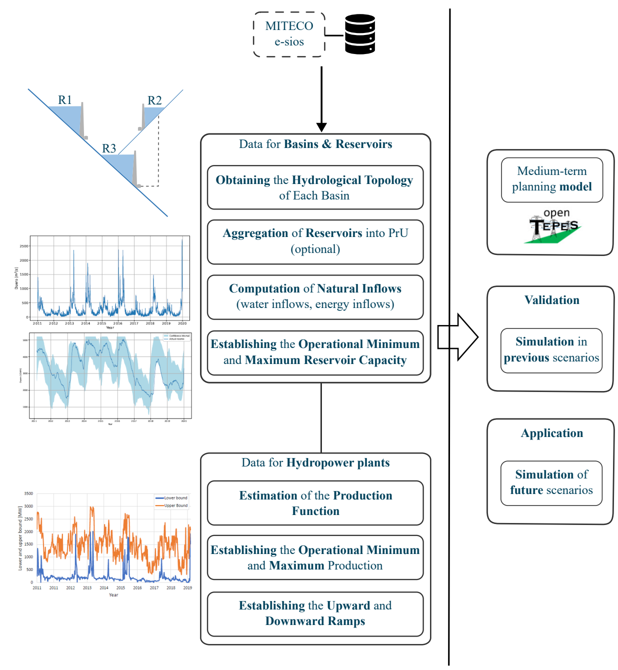

Figure 1: Methodology for Hydropower Modeling in Medium-Term Planning

The following methods are described in the following subsections:

- Auxiliary Linear Optimization Model
- Fourier Series Filtering and Linear Regression Models
- Load-Duration curve based on K-means algorithm

3.1. Basins & Reservoirs

A basin is a geographical area where precipitation collects and drains into a common outlet, such as a river, lake, reservoir, or a combination of these. Within this context, reservoirs play a critical role by storing water as the primary energy source for hydropower generation. Therefore, it is crucial to establish the basin's topology, identify reservoir aggregations (if required to simplify model complexity), quantify the natural inflows to physical and equivalent reservoirs (derived from these aggregations), and determine their operational minimum and maximum capacities. To address these considerations, we present a set of methods described below.

Obtaining the Hydrological Topology of Each Basin

The basin's topology includes a network of reservoirs, production units, and pumping units connected by rivers, streams, weirs, and canals. Accurately modeling hydropower production requires obtaining this topology. For this task, we employ GIS maps developed by MITECO (Section 2).

Aggregation of Reservoirs into PrU

Various criteria can be employed for this aggregation, as outlined in Härtel and Korpås (2017). However, using pre-established market groupings is advantageous due to the availability of additional data for constructing the equivalent model.

In many instances, when an administration allocates exploitation rights over reservoirs to various companies within a basin, it may assign the exploitation rights for the entire basin or a majority of it to a single company. This consolidation enables more efficient administration of hydro resources, owing to the significant coupling interactions among the management of different reservoirs within the same basin. This aggregation is commonly known as a Programming Unit (PrU).

Aggregating hydropower plants into PrU results in a loss of detailed information about the individual power generation of each plant over time, making only the total aggregate production available. Despite the inherent loss of information, aggregating hydropower plants into PrU is the best way to build an equivalent hydropower plant since it can capitalize on the existing market records, such as power production and electrical consumption.

Computation of Natural Inflows

The natural inflows to a reservoir refer to the water entering the reservoir, excluding any water released from other hydropower plants and spillages from upstream reservoirs.

The described procedure involves two main steps:

- **Calculating Natural Water Inflows**: Calculate the natural water inflows to each reservoir (physical representation).
- **Calculating Natural Energy Inflows**: Convert these natural water inflows into natural energy inflows for an equivalent aggregated energy hydropower plant.

To estimate the **natural water inflows** from open data in Spain, we employed an auxiliary linear model that considers the topology of each PrU, the production function of every hydropower plant, reserve levels, pumping production, and other publicly available parameters discussed in previous sections.

The mathematical model is constrained by a modified equation of the water balance that includes information regarding pumping consumptions and the topology of each PrU. The linear model is explained below.

According to figure 2, the water balance for each reservoir r is given by the equation 1. This equation captures the daily dynamics of a hydroelectric reservoir's reserves, accounting for the balance between natural processes, energy production, inter-reservoir coordination, and water usage demands.

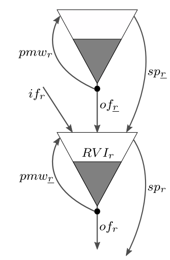

Figure 2: Reservoir modeling

$$\begin{aligned} RV_r = \quad & RVI_r + if_r - of_r - pmw_r - sp_r \\ & + of_{to(r)} + pmw_{to(r)} + sp_{to(r)} - ir_r + aif_r \end{aligned} \tag{1}$$

If we consider $OUT_r = of_r + pmw_r$ and the water outflows from each hydropower plant end up in the same reservoir downstream, then:

$$of_{to(r)} = \sum_{\underline{r}\in\mathcal{R}_r} of_{\underline{r}} \tag{2}$$

If we add and subtract $\sum_{\underline{r}\in\mathcal{R}_r} pmw_{\underline{r}}$ to equation 2, we obtain

$$\begin{aligned} of_{to(r)} &= \sum_{\underline{r}\in\mathcal{R}_r} \left(of_{\underline{r}} + pmw_{\underline{r}} \right) - \sum_{\underline{r}\in\mathcal{R}_r} pmw_{\underline{r}} \\ &= \sum_{\underline{r}\in\mathcal{R}_r} OUT_{\underline{r}} - \sum_{\underline{r}\in\mathcal{R}_r} pmw_{\underline{r}} \end{aligned} \tag{3}$$

Incorporating the relation 3 into equation 1 results in equation 4b, which serves as a constraint within the optimization model to calculate the natural water inflows of each reservoir.

Considering the aforementioned analysis, we formulate the optimization model 4 to calculate the natural inflows of each reservoir.

$$\min \quad \sum_r \left(ir_r + PN.sp_r + PN2.aif_r \right) \tag{4a}$$

s.t.

$$\begin{aligned} RV_r = \quad & RVI_r - OUT_r + \sum_{\underline{r}\in\mathcal{R}_r} OUT_{\underline{r}} + if_r - sp_r - ir_r + \\ & + \sum_{\underline{r}\in\mathcal{R}_r} sp_{\underline{r}} - \sum_{\underline{r}\in\mathcal{R}_r} pmw_{\underline{r}} + pmw_{to(r)} + aif_r, \quad \forall (r) \end{aligned} \tag{4b}$$

$$PU = 24 \sum_p pm_p \tag{4c}$$

$$pmw_r = 24 \sum_{p\in\mathcal{R}_p} EFF_p pm_p / EC_p, \quad \forall (r) \tag{4d}$$

$$pmw_{to(r)} = 24 \sum_{p\in\mathcal{P}_r} EFF_p pm_p / EC_p, \quad \forall (r) \tag{4e}$$

$$\underline{PM}_p \leq pm_p \leq \overline{PM}_p, \quad \forall (p) \tag{4f}$$

The model aims to find a feasible solution by prioritizing slack variables in the objective function. It first utilizes irrigation operations, then spillages, and finally, artificial water inflows to prevent infeasibilities. This sequence reflects the logical order of these variables according to real-world operational practices. The model is designed to operate on a daily resolution consistent with the publicly available parameters.

The conversion from natural water inflows to **natural energy inflows** follows the concept described in Arvanitidits and Rosing (1970). We obtain the natural energy inflows by applying an aggregated conversion factor to each reservoir's natural water inflows and summing them (Figure 3).

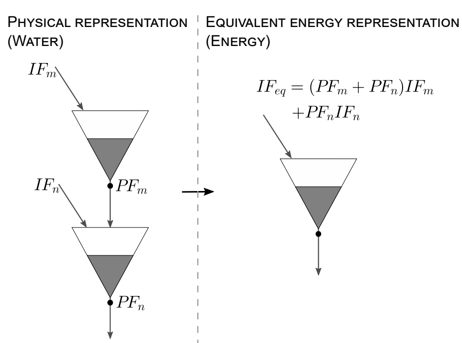

Figure 3: Conversion from Physical Representation to Equivalent Energy hydropower Plant

Establishing the Operational Minimum and Maximum Reservoir Capacity

Hydropower reservoirs require establishing storage limits that vary depending on the season of the year and the corresponding water inflow levels. Operational requirements for reservoirs differ significantly between winter and summer, with winter focusing on water storage for electricity production and summer prioritizing other uses.

A common approach to defining these limits involves utilizing historical minimum and maximum values for each month. However, this method may result in overly broad ranges, leading to an excessively flexible model that does not accurately reflect real system constraints.

We propose a methodology that involves the following steps:

1. Standardize and normalize the data (reservoir and water inflows) using equations 5a and 5b correspondingly.

$$\overline{\mathbf{Y}}(t) = \left[\frac{y_1(t) - \mu_1}{\sigma_1}, \ldots, \frac{y_i(t) - \mu_i}{\sigma_i}\right] \tag{5a}$$

$$T\left(y_i(t)\right) = \frac{y_i^{\lambda}(t) - 1}{\lambda} \tag{5b}$$

where $y_i(t)$ represents each time series and λ is the parametric value of the Box-Cox transformation.

2. The next step involves filtering the explanatory variables, specifically the water inflows. This step is crucial for identifying seasonal trends and correlations with the variables to be predicted (Hassani, Mahmoudvand and Yarmohammadi (2010), Meng, Wang, Guo and Ding (2023)). In this study, we employ a low-pass filter applied to the water inflows, guided by the observation that the reservoir level exhibits a smooth behavior. Generally, system operators do not make daily adjustments to the operation of large reservoirs; instead, the operation of these reservoirs relies on the overall behavior of water inflows within a given week or month. Mathematically, this implies that the reservoir level does not contain high-frequency components.
   Various methods are employed to execute this filtering process, including moving averages at different frames Aradhye, Rao and Mastan Mohammed (2019), Fourier analysis Oppenheim, Willsky and Nawab (1997), and wavelet transform Joo and Kim (2015). In this study, we specifically focus on the first two methods. These techniques are numerically compared using R^2, $RMSE$, and MAE scores based on the linear regression model. The details of these filtering methods are outlined below:

   Moving Averages at Different Frames: The moving averages method involves calculating averages to smoothen the explanatory variables. We employed weekly, two-week, and monthly frames for the moving averages.

   Filtering Based on Fourier Analysis: In this method, we analyze the frequency spectrum of the variable to be predicted (reserve level). The resulting spectrum is then compared with the explanatory variables (water inflows). This step is crucial in determining whether the predictor variables need to be filtered, thereby averting potential issues associated with fitting the parametric model in the presence of high-frequency components.

3. A linear regression model is utilized in the modeling phase, but it is noteworthy that other alternative models, including neural networks, can be explored for this purpose. In this study, we specifically use a linear regression model with Ridge regularization. The weight for controlling the regularization strength is a hyperparameter calculated through a cross-validation method over a grid of potential weights. The R^2 determination score is then employed to assess the predictability of each weight, aiding in the selection of the best one.

4. Calculate the standard deviation of the error between the predicted value and the actual variable. This standard deviation is used to define the range of reservoir levels. In this study, we employ 2 standard deviations, as indicated in Equation 6.

$$rv_r \in \left[\hat{rv}_r - 2\sigma_r, \hat{rv}_r + 2\sigma_r\right] = \left[LB_r, UB_r\right], \quad \forall (r) \tag{6}$$

In this context, rv_r represents the reservoir level obtained from the optimization model, $\hat{rv}_r$ denotes the predicted value from the parametric model, and σ_r indicates the standard deviation of the errors between historical and expected values.

In this analysis, data is considered at a weekly resolution, calculated as the average of values within each week. This method applies to both physical and equivalent energy representations. For the physical representation, the analysis is performed on each reservoir using MITECO data. For the equivalent energy representation, MITECO data is first converted into equivalent energy, as outlined in Figure 3, and then the procedure is applied.

3.2. Hydropower Plants

This section outlines the process for determining the parameters of hydropower plants in both the physical and energy representations. This study considers the production function, the minimum and maximum production levels, and the upward and downward ramps for hydropower plants.

Estimation of the Production Function

The hydropower production function describes the relationship between energy production and both the water release and net head in the reservoir (Cerisola et al., 2012). While obtaining the parameters of this function publicly can be challenging, data related to reservoir levels, both in terms of water and energy units, are commonly available. These data serve as an approximation of production based on the release and net head. The production function is relevant for estimating water inflows and converting physical reservoirs into equivalent energy hydropower plants.

Establishing the Operational Minimum and Maximum Hydropower Plants Production

A hydropower plant's minimum and maximum power production are well defined by its technical characteristics. However, for mid-term deterministic planning models, it is important to adjust these bounds to address the inherent limitations of such models in capturing uncertainties. These adjustments ensure that the planning models more accurately reflect the real-world power production variability due to seasonal water availability and operational constraints.

Determining the minimum and maximum power production for an equivalent hydropower plant requires a more nuanced approach than merely summing the technical characteristics of individual plants. Such a simplistic method would result in an equivalent plant with excessive flexibility, as it unrealistically allows for the simultaneous maximum production of all the plants.

In real-world operations, the power production of hydropower plants is closely tied to seasonal variations, with higher production during wet seasons and lower output during dry periods. Various methods can be applied to account for these fluctuations effectively:

- **Minimum and Maximum values:** This approach periodically determines minimum and maximum power production, such as daily or weekly.
- **Quantiles:** Defining quantiles, such as the 1st and 99th percentiles, weekly or by another period, helps to represent the minimum and maximum power production range. This approach aims to exclude outlier situations from the past that may not be representative of future expectations.
- **Duration-Curve Method:** This technique involves constructing duration curves to analyze the distribution of power production over time. Examining this duration curve makes it possible to determine the minimum and maximum power production levels closely reflecting actual operational conditions. In this work, the duration curve is obtained through clustering methods. Figure 4 describes the procedure to obtain the production bounds.

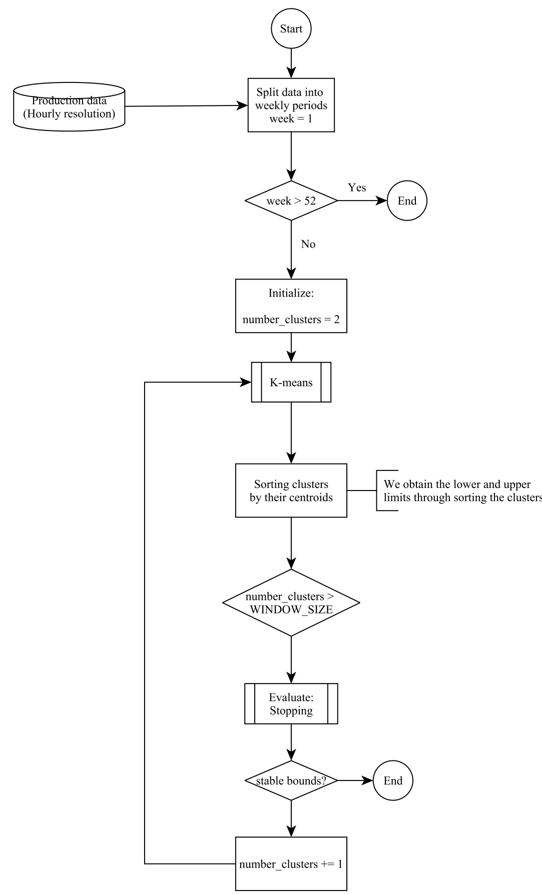

Figure 4: Method to obtain the lower and upper production bounds

In this study, we establish the process as stable when the lower and upper bounds do not deviate by more than 5% from the moving average of the last three iterations, as depicted in Equation 7.

$$
\begin{aligned}
\text{stop} &= \begin{cases} UB_n \in \left[0.95\hat{UB}_n, 1.05\hat{UB}_n\right] & \text{and} \\ LB_n \in \left[0.95\hat{LB}_n, 1.05\hat{LB}_n\right] \end{cases} \\
\hat{UB}_n &= \frac{UB_{n-1} + UB_{n-2} + UB_{n-3}}{3} \\
\hat{LB}_n &= \frac{LB_{n-1} + LB_{n-2} + LB_{n-3}}{3}
\end{aligned}
\tag{7}
$$

where, UB_n and LB_n represent the upper bound and lower bound with n clusters.

Establishing the Upward and Downward Ramps

Similar to the techniques used to determine the minimum and maximum power expected from a PrU at a specific period, the maximum ramps were obtained through a statistical analysis of historical data.

For each PrU, data regarding the increase or decrease in power supplied from one hour to the next has been calculated across all available data. These data points have been grouped into a predetermined number of clusters per PrU, from the largest increase to the largest decrease. In this computation, we employ the same clustering method for hourly production. The clusters representing the largest increase and the largest decrease are dropped from the analysis. This filtering step excludes exceptional operating conditions and maintains only the operational upward and downward ramps.

4. Case study and results

This section presents the application of the proposed methods through a comprehensive case study in Spain. Section 4.1 focuses on the extraction of the hydropower parameters mentioned in section 3. Section 4.2 addresses the validation process, where the parameters are tested by configuring a hydrothermal optimization model and comparing its results against historical data to assess accuracy and reliability. Finally, Section 4.3 demonstrates the application of the validated parameters within the context of medium-term planning by executing a case study for 2025, showing the model's practical utility and performance. The Spanish energy system was chosen as the case study for evaluating the performance of the hydropower modeling proposed in this paper.

4.1. Hydropower Parameter Extraction

4.1.1. Basins & Reservoirs

Topology of Basins

We have obtained the topological representation of the Duero, Tajo, and Sil basins, which contribute to approximately 56% of the reservoir-hydro production within the Spanish energy system (Figure 5).

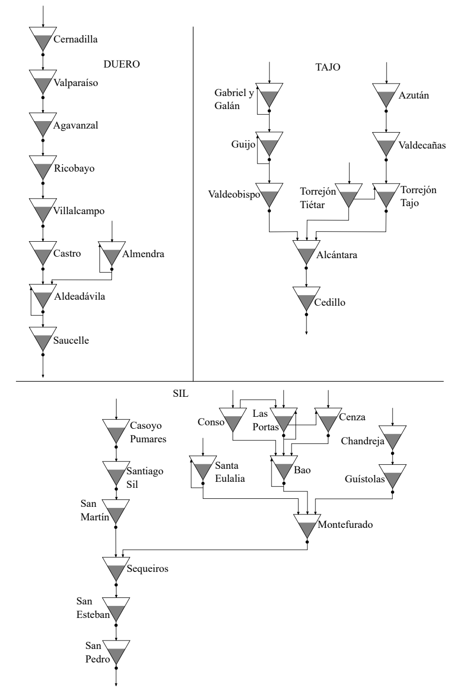

Figure 5: Topology of Duero, Tajo, and Sil basins

Water inflows / Reserve	DUER	TAJO	SIL
Annual	0.28	0.32	0.35
Annual + Filtering	0.39	0.37	0.37
Jan + Filtering	0.77	0.31	0.61
Feb + Filtering	0.74	0.36	0.68
Mar + Filtering	0.64	0.66	0.72
Apr + Filtering	0.66	0.82	0.77
May + Filtering	0.63	0.84	0.71
Jun + Filtering	0.65	0.87	0.63
Jul + Filtering	0.68	0.89	0.49
Aug + Filtering	0.76	0.88	0.37
Sep + Filtering	0.70	0.76	0.34
Oct + Filtering	0.51	0.56	0.27
Nov + Filtering	0.42	0.16	0.30
Dec + Filtering	0.50	0.05	0.39

Table 1
Correlation coefficient between reserve and water inflows

Natural Inflows

We computed the natural inflows for the Duero, Tajo, and Sil basins, represented by the PrUs DUER, TAJO, and SIL, respectively. Following the methodology outlined in Section 3.1, we generated representations for both natural water inflows and natural energy inflows.

Figure 6 and Figure 7 show the natural water and energy inflows for each basin.

Operational Minimum and Maximum Reservoir Capacity

Section 3.1 described the methodology for deriving the curve-rule operation. To improve the predictability of the linear model, we performed a monthly analysis after applying filtering methods. The correlation results are presented in Table 1.

Based on the correlation analysis results in Table 1, we developed linear regression models for each month, correlating reserve levels with water inflows. The analysis includes a comparison of various filtering methods, including moving average smoothing and Fourier analysis. Model performance was assessed using R^2, RMSE, and MAE scores, as shown in Table 2 for the PrU DUER under different filtering methods. These findings led us to apply Series-Fourier filtering to the predictor variables (water inflows and their lags) to improve the linear model's accuracy between reserve levels and water inflows.

Figure 8 shows the cross-correlation results for the PrU DUER between the reserve and the water inflows. It is

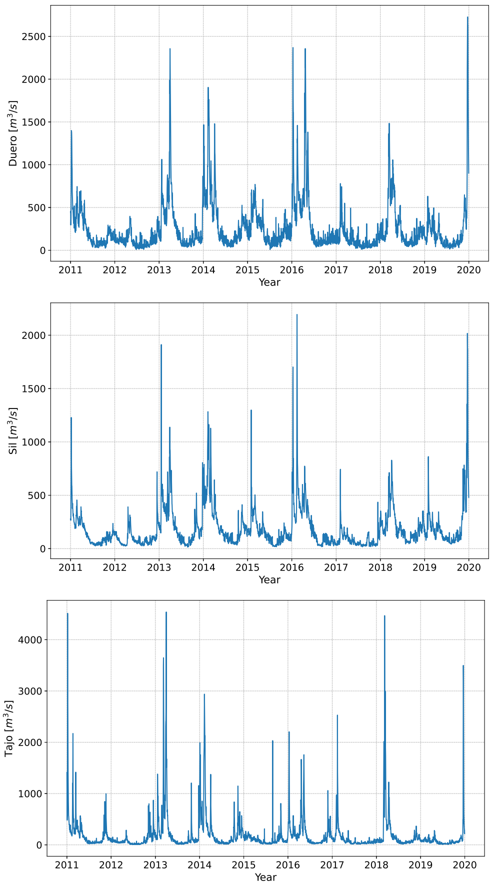

Figure 6: Water Inflows of Production Units $[\mathrm{m}^3/\mathrm{s}]$

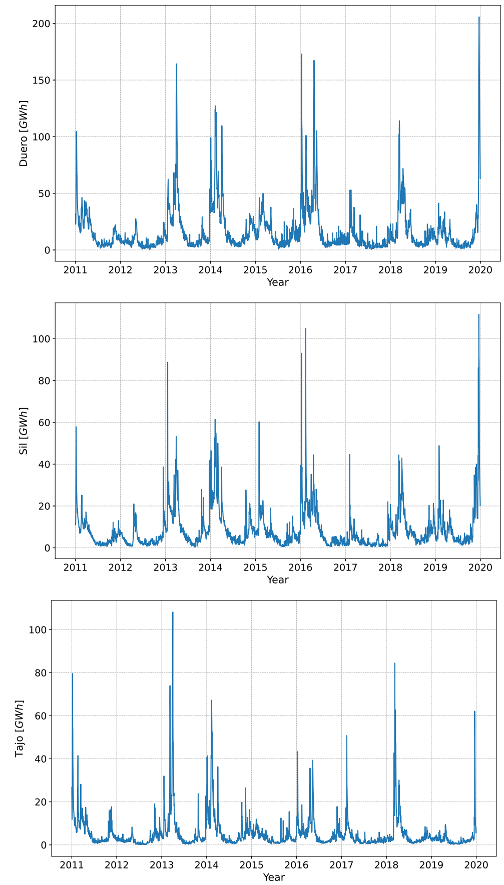

Figure 7: Energy inflows of Equivalent Hydropower Units [GWh]

Table 2
Comparison of Filtering Methods for Deriving the Linear Regression Model Between Reserve Level and Water Inflows for PrU DUER

Method	R^2	$RMSE$	MAE
Moving average. Weekly	0.65	612	520
Moving average. Two-weeks	0.64	616	521
Monthly	0.63	629	526
Fourier series	0.76	511	420

important to note that the predictor variables in this analysis are the monthly average water inflows. Although the water inflows are measured at a daily resolution, we aim to predict each reserve value based on previous monthly trends of the water inflows.

Following the procedure described in Section 3.1, the limits for PrU DUER, TAJO, and SIL are shown in Figure 9. This interval predicts the reserve levels with 95% confidence across different water inflow scenarios (dry, average, and wet).

4.1.2. Hydropower plants

Operational Minimum and Maximum Hydropower Plants Production

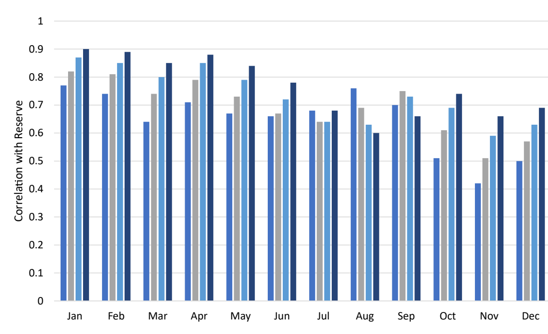

Figure 8: Cross correlation between PrU DUER reserve and the monthly trends of the water inflows at different lags (1m = 4 weeks)

To illustrate the effectiveness of the analytical methods described in Section 3.2, the PrU DUER, one of Spain's most significant, was selected. This basin includes fourteen hydropower plants with a combined maximum design power capacity of 3,460 MW. Theoretically, these turbines can be committed without producing energy, resulting in a minimum power of 0 MW.

The study analyzed hourly power production data from 2011 to 2019. Table 3 presents annual statistical measures, including the technical minimum and maximum power

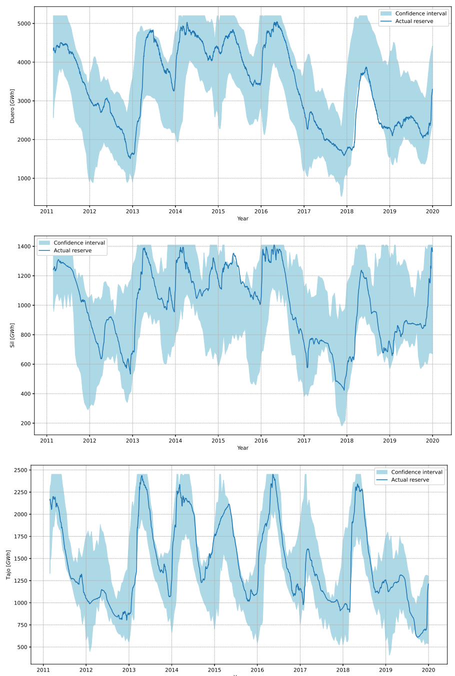

Figure 9: Operational Minimum and Maximum Reservoir Capacity

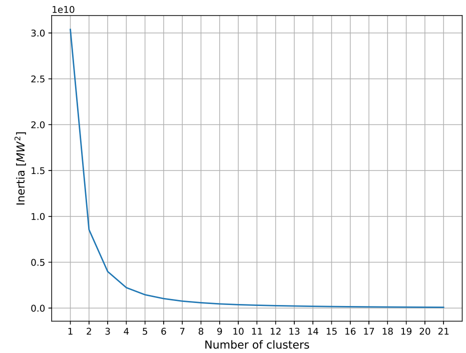

Figure 10: Elbow rule

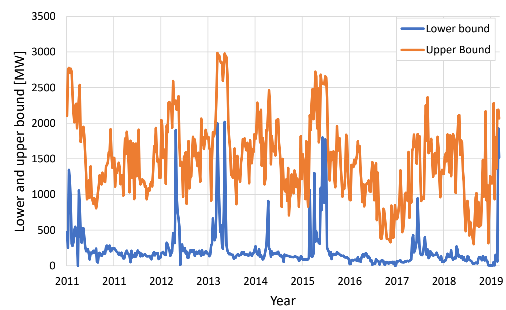

Figure 11: Weekly minimum and maximum power for PrU DUER. Dynamic Load-curve duration

Table 3
Annual statistical measures. PrU DUER
* 13 clusters

Method	$P_{min}\,[MW]$	$P_{max}\,[MW]$
Technical characteristics	0	3460
Historic min and max	0	3201
1st and 99th percentile	31	2643
Elbow rule*	97	2813
Load-curve duration	63	2867

production, as well as the minimum and maximum values derived from the 1st and 99th percentiles and load-curve duration analysis. For comparison, two duration curves were constructed using two methods to select the number of clusters: The Elbow rule and the method proposed in this study.

According to the Elbow rule, the optimal number of clusters is thirteen. By using our proposed stopping criterion, the clustering process requires twenty-one clusters. Table 3 compares the Elbow rule and our method results. While both yield similar values for the variable limits, our method focuses on stabilizing the bounds, which are neither as wide as the technical characteristics or the historical minimum and maximum of the time series nor as narrow as the 1st and 99th percentiles.

In this work, we implemented the proposed method weekly throughout the time series. Figure 11 shows the results for the PrU DUER.

Figure 12 illustrates the stabilization of the power production limits for the PrU DUER, following the procedure outlined in section 3.2. In this instance, the number of clusters in our methodology is interpreted as the number of bins in the load-curve duration.

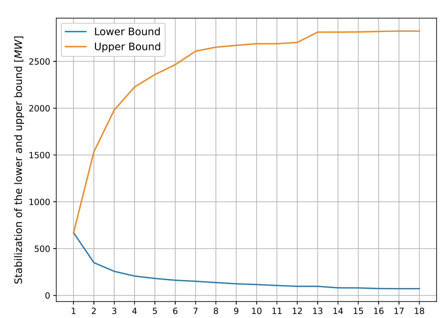

Figure 12: Load-Curve Duration Analysis: Relationship between bounds stabilization and the number of clusters used. Whole time series

Upward and Downward Ramps

Following a similar procedure to calculate the minimum and maximum power production, we determined the operational upward and downward ramps for equivalent hydropower plants. The resulting values are presented in table 4.

Table 4
Upward and Downward Ramps

Hydropower plant	Downward $[MW/h]$	Upward $[MW/h]$
DUER	745	790
SIL	300	305
TAJO	565	570

4.2. Validation: Year 2019

The parameters required for the optimization model are configured using the results from the analytical methods outlined in section 3. To assess the quality of the modeling, we configured the openTEPES model Ramos, Alvarez and Lumbreras (2022), a standard unit commitment model, to evaluate the operation for the year 2019. Subsequently, the results of the hydropower production and reserve levels are compared with the actual values for that year.

The case study highlights the following characteristics:

- Detailed modeling of three basins: Duero and Sil as equivalent energy plants, and Tajo represented by its 9 cascaded water reservoirs.
- 35 PrU units with capacities exceeding 50 MW.
- 51 generation plants, including combined cycle gas turbines (CCGT) and coal-fired units.
- 7 nuclear plants.
- Renewable generation, including thermal solar, photovoltaic solar, and wind production.
- Other generation sources, including biogas, biomass, geothermal, and cogeneration.
- Consideration of 8736 hours in the optimization model.

It is essential to note some approximations made in this study, which are common when working with open data in an electrical system. First, the variable costs of thermal plants were considered every month rather than a daily one. Second, solar and wind production plants were aggregated into one representative programming unit for each technology. Maintenance scheduling and forced outages were also modeled using an Equivalent Forced Outage Rate (EFOR). Finally, a deterministic optimization model was employed.

Figure 13 compares the performance of various generation technologies based on the simulation results of the Spanish electricity system for 2019 and the actual observed values, Red Eléctrica de España (2019). The maximum discrepancies, approximately 0.3%, are primarily associated with hydropower and CCGT+Coal production.

Figure 14 depicts the water management in two significant reservoirs within the Tajo basin, offering a closer look at how water levels are managed.

Figure 15 showcases the energy management in the Duero and Sil basins, the other two critical basins in the Spanish system, providing additional insights into energy production and utilization dynamics.

Figure 14 and Figure 15 demonstrate a strong correlation in the trends of reservoir levels between the results of

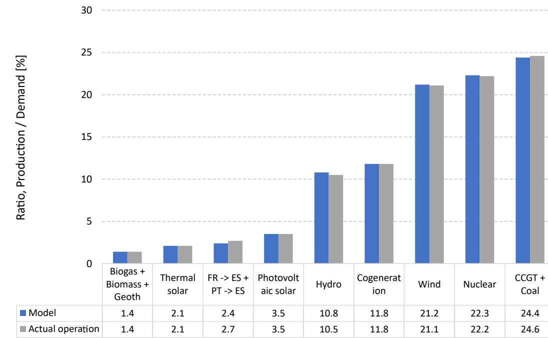

Figure 13: Comparison of technology production between the simulation results and the actual 2019 operation

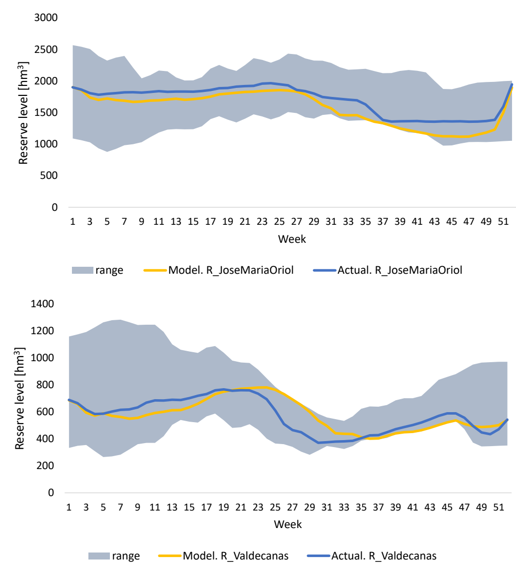

Figure 14: Reserve level. Real reservoirs. 2019 operation

the model (openTEPES) and the observed behavior. Table 5 shows the R^2 measure for this case study, labeled as "Benchmark."

A deviation in the PrU SIL is evident around week thirty-seven. This discrepancy arises from the specific conditions of the water inflows in 2019. During this week, water inflows increased. In actual operation, the decision was made to store water due to the uncertainty of future inflows. However, the model used in this study is deterministic, allowing for better water use since it "knows" the future inflows. It is important to note that the imposed limits on reservoir levels restrict the model's flexibility in decision-making, resulting in behaviors more similar to real operations.

4.2.1. *What-If Analysis*

We conducted a sensitivity analysis to examine the impact of replacing the parameters calculated in Section 4.1 (proposed methods) with design values representing the technical capabilities of hydropower plants and reservoirs. These design values are less restrictive within the optimization model than the computed parameters, which are adjusted to

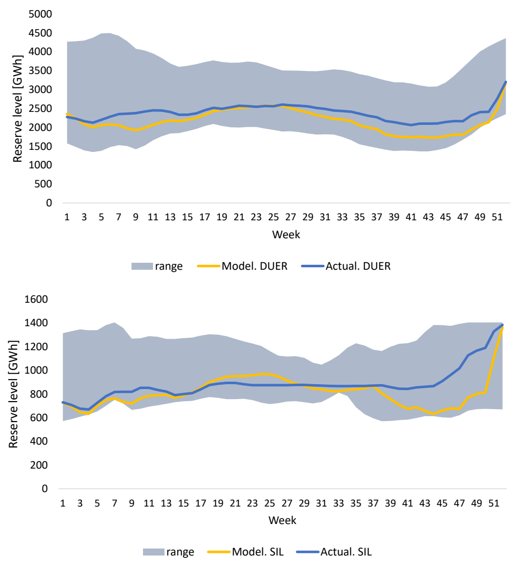

Figure 15: Reserve level. Equivalent reservoirs. 2019 operation

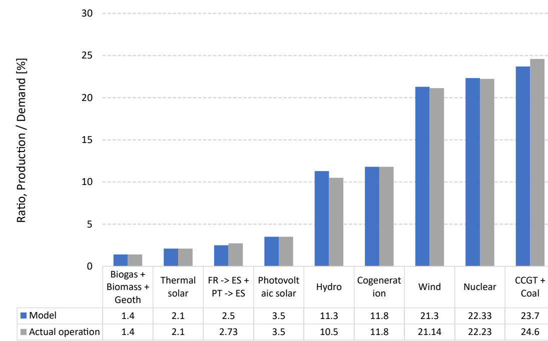

Figure 16: Comparison of Technology Production: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants Instead of Operational Limits

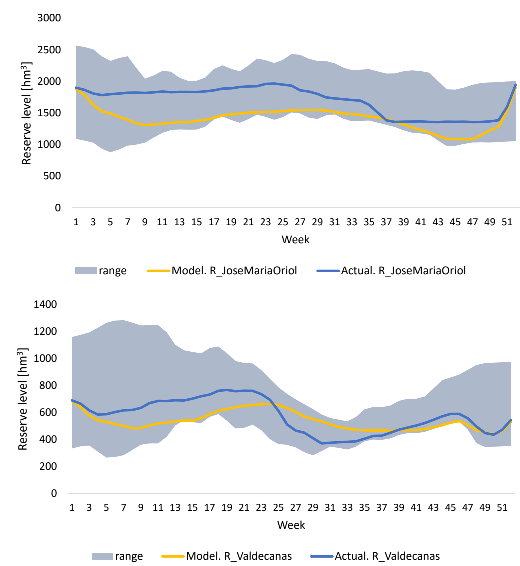

Figure 17: Comparison of Reserve Levels in **Real** Reservoirs: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants Instead of Operational Limits

reflect practical operational constraints. This analysis focused on varying reservoir limits, production limits, and water inflows to evaluate their influence on the results. The 2019 case study was performed with three different variations:

- *Sensitivity 1, S1*: In this case, the design values for hydropower plant production limits are applied consistently throughout the year rather than using the operational minimum and maximum power production values calculated in Section 4.1.
- *Sensitivity 2, S2*: This case builds upon the previous sensitivity (S1). Additionally, reservoir design values are applied throughout the year, replacing the operational minimum and maximum reservoir limits calculated in Section 4.1.
- *Sensitivity 3, S3*: This case combines the design values for hydropower plant production and reservoir limits from the previous cases. Furthermore, it incorporates historical power production data as inflows, replacing the natural inflows calculated in Section 4.1.

S1. Design Values for Hydropower Plants Replacing Operational Limits

Figure 16 shows a higher hydropower production than the scenario with weekly production constraints. This increase in hydropower output leads to a reduction in CCGT+Coal production to meet electrical demand. Although the overall annual difference between these scenarios may appear minimal, Figures 17 and 18 highlight significant variations in reservoir operations. Greater operational flexibility for hydropower results in different decision-making across different year seasons.

S2. S1 + Design Values for Reservoir Limits Replacing Operational Limits

Figure 19 shows that the overall annual production of hydropower and CCGT+Coal technologies remains the same compared to the previous sensitivity (S1). However, Figures 20 and 21 reveal that reservoir operations differ significantly from actual operations. For instance, the PrU SIL operates flexibly, reducing reserve levels to zero. Such operations are uncommon in real systems due to the inherent uncertainties in water inflows and other operating conditions.

In this sensitivity, the design reservoir limits replaced the operational limits and were applied consistently throughout the year. The grey zones in Figures 20 and 21 are only indicative references. Employing design values provides the model with greater flexibility in decision-making, resulting in a more cost-efficient solution; however, it also leads to more significant deviations from actual operations. This discrepancy arises primarily because the deterministic model assumes perfect knowledge of future water inflows, a condition that does not hold in real-world systems operation. This fact highlights the importance of configuring operational limits for reservoirs to balance the model's decision-making

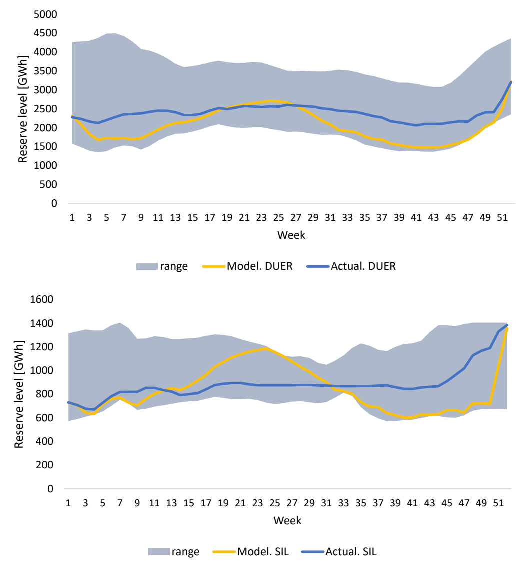

Figure 18: Comparison of Reserve Levels in **Equivalent** Reservoirs: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants Instead of Operational Limits

flexibility, ensuring it does not overly exploit perfect foresight and produce unrealistic operations.

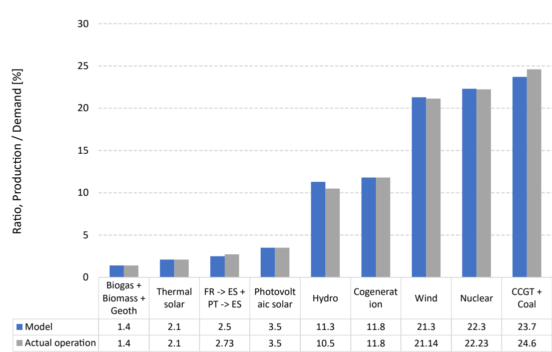

Figure 19: Comparison of Technology Production: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants and Reservoirs Instead of Operational Limits

S3. S2 + Production as Inflows Replacing Natural Water Inflows

Figure 22 depicts that the overall annual energy production exhibits the same behavior as in Figure 13 (Case "Benchmark"). However, two important points must be kept in mind:

- The operation of the reservoirs differs from the actual operations in 2019. Although the real production for 2019 was used as inflows to the PrU, the model does not replicate the operations for that year. Instead, the model re-optimizes the operations for 2019. This fact occurs because the model is deterministic and optimizes the system with “inflows" that are not natural inflows. Figures 23 and 24 illustrate these differences.

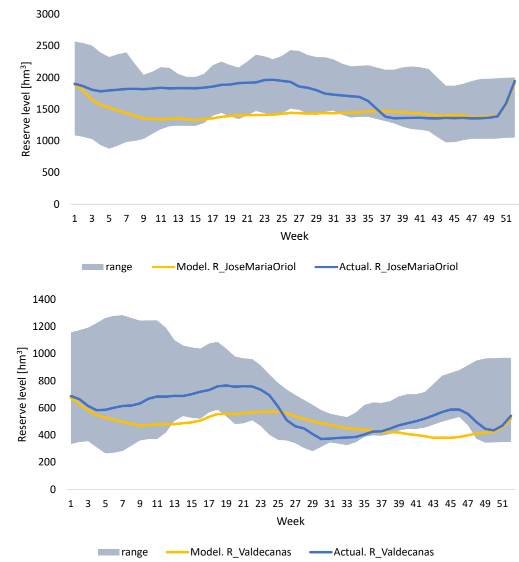

Figure 20: Comparison of Reserve Levels in **Real** Reservoirs: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants and Reservoirs Instead of Operational Limits

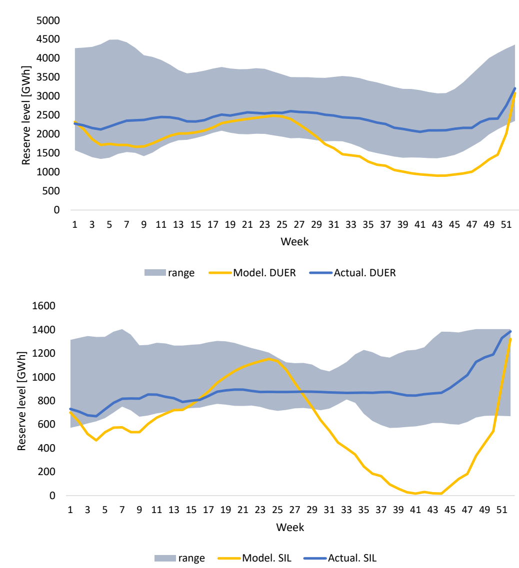

Figure 21: Comparison of Reserve Levels in **Equivalent** Reservoirs: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants and Reservoirs Instead of Operational Limits

- Using these productions as inflows for future scenarios is an incorrect assumption. While optimal productions are obtained under specific conditions of a climate year, it is important to remember that these conditions will not be the same in the future (e.g., increased renewable generation compared to past conditions). Considering productions as inflows results in the loss of natural inflow management and, therefore, yields incorrect results.

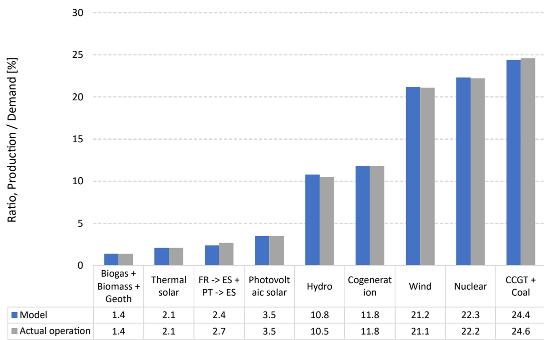

Figure 22: Comparison of Technology Production: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants and Reservoirs Instead of Operational Limits, and Actual Production as Inflows

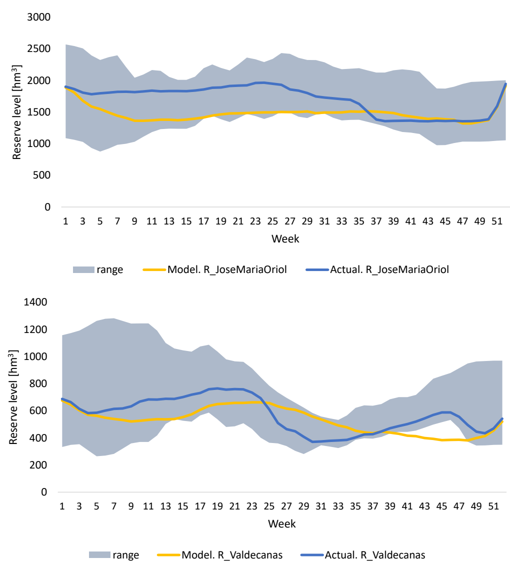

Figure 23: Comparison of Reserve Levels in **Real** Reservoirs: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants and Reservoirs Instead of Operational Limits, and Actual Production as Inflows

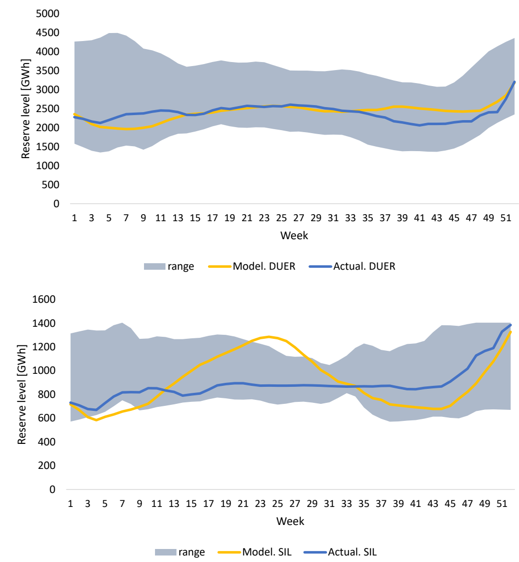

Figure 24: Comparison of Reserve Levels in **Equivalent** Reservoirs: Model Results vs. Actual Operation in 2019, Considering Design Values for Hydropower Plants and Reservoirs Instead of Operational Limits, and Actual Production as Inflows

Table 5
Comparison of different parameter combinations for hydropower plants: R^2 measure between model results and actual operation data for 2019

Reservoir	Benchmark	S1	S2	S3
José María Oriol	0.98	0.71	0.16	0.36
Valdecañas	0.76	0.64	0.52	0.55
DUER	0.91	0.87	0.77	0.48
SIL	0.59	0.30	0.15	0.46

Table 5 summarizes the different sensitivity results. This table presents the R^2 measure, comparing the results regarding technology production over the year and the reserve levels obtained through the model with the actual operation data for 2019. The column labeled “Benchmark" refers to the case study using the parameters calculated through the methods proposed in this paper. The columns labeled “S1," “S2," and “S3" correspond to the previously mentioned sensitivity analyses.

The results presented in Table 5 highlight the effectiveness of the methods proposed in this study, particularly in the “Benchmark" case, where the metrics consistently approach 1, indicating better alignment with real-world operations. In contrast, the sensitivity analyses (S1, S2, S3), which replace the operational parameters derived using the proposed methods with design values, reveal notably lower performance metrics. In each sensitivity case, the operation suggested by the optimization model diverges from actual operations due to the increased flexibility in the decision-making process. Specifically, the use of design values for hydropower plant production and reservoir limits, or the use of actual production as water/energy inflows, introduces a level of flexibility that is more theoretical than practical, resulting in model outputs less representative of real-world conditions.

4.3. Application: Year 2025

In this section, we have simulated the Spanish energy system for the year 2025. In addition to the data referenced above, we have compiled other essential information required for the setup of the case study:

- **Demand:** The National Resource Adequacy Assessment (NRAA) report REE, which extends the European Resource Adequacy Assessment (ERAA), projects a peak demand of 46.45 GW and an annual energy consumption of 258.68 TWh for 2025. In this case study, hourly load distribution factors from 2023 are scaled to match the 2025 peak demand, as illustrated in Figure 25.
- **Wind and solar production:** The "PEMMDB Generation" file, referenced in ENTSOE, provides projected installed

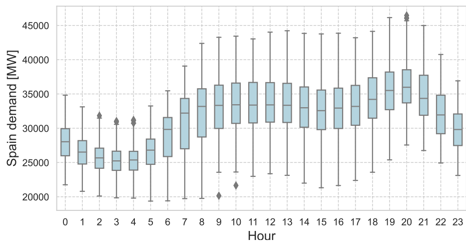

Figure 25: Spain demand expected for 2025

capacities for different generation technologies. For 2025, it estimates 34,817 MW for wind and 33,698 MW for photovoltaic solar. The hourly load factors align with the simulated climate year, and in this study, we used 2019's climate conditions to model 2025.

For this analysis, we have incorporated the trends in production technologies and the behavior of the reserve levels of the primary basins in the Spanish system. We compared these results with the 2019 operation in relative values to the demand.

Figure 26 depicts shifts in energy production trends. Solar and wind energy generation increase, while CCGT+Coal production drops to meet demand. On the other hand, there is a significant rise in hydropower, mainly due to increased pumping operations. Specifically, hydropower production through pumping rises from 1642 GWh in 2019 to 8161 GWh in 2025. Excluding pumping, hydropower generation shows a minor change from 26092 GWh in 2019 to 26263 GWh in 2025.

The figure also compares production technologies with 2023 data (Red Eléctrica de España (2023)), offering a more recent perspective. However, it's worth noting that hydrological conditions in Spain differed between 2023 and 2019, warranting caution in interpreting the overall values.

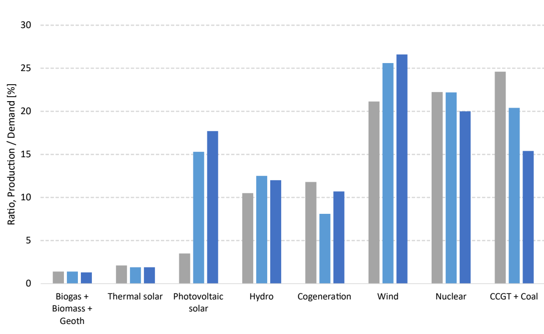

Figure 26: Comparison of technology production between the expected operation for 2025 and the real operation for 2019 and 2023

When examining the trends in reservoir levels, we note a similarity in reservoir operations between 2019 and 2025. In the main reservoirs, given the comparable climate conditions to those of 2019, there are no observed spillage conditions or requirements for water inflows. A notable increase in pumping operations is observed, throwing from 1642 GWh in 2019 to 8161 GWh in 2025. The integration of renewable resources mainly drives this fact. Figure 27 illustrates the operations of José María Oriol y Valdecañas reservoirs for both 2019 and 2025, while Figure 28 displays the behavior of the equivalent PrU DUER and Sil for the same years.

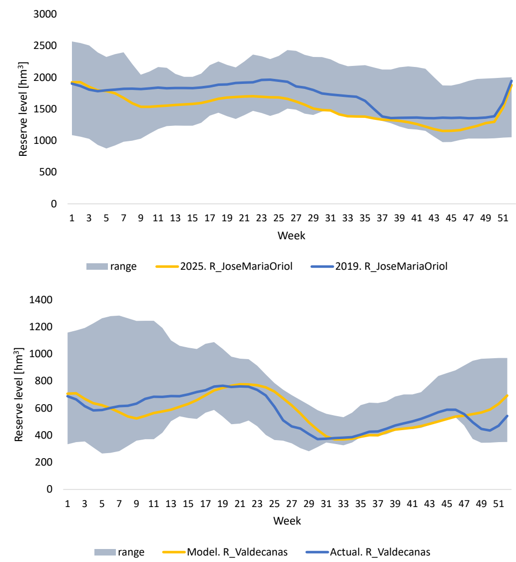

Figure 27: Reserve level. Real reservoirs. 2025 vs. 2019 operation

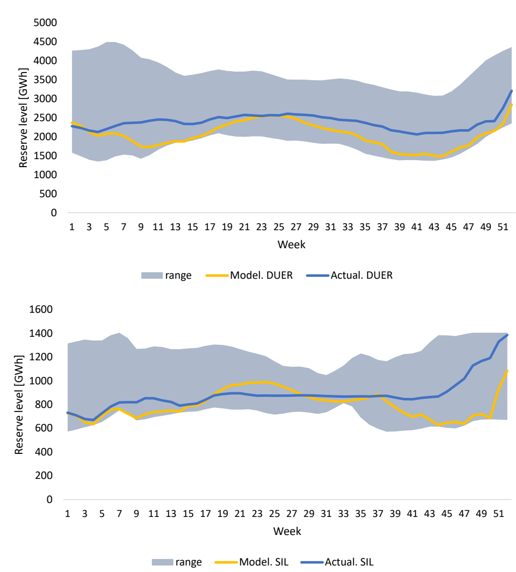

Figure 28: Reserve level. Equivalent reservoirs. 2025 vs. 2019 operation

5. Discussion

5.1. Operational Minimum and Maximum Power Production

Analyzing minimum and maximum power production is crucial for configuring planning models and even more for studying the increasing integration of renewable energy. Summing the individual hydropower plants' minimum and maximum production capacities may result in overly flexible equivalent plants, not reflecting real system constraints (Table 3).

Water inflows, strongly correlated with power production, serve as excellent explanatory variables. Modeling minimum and maximum power production based on water inflows provides a robust framework for parameter configuration.

5.2. Operational Minimum and Maximum Reservoir Capacity

Linear regression models are beneficial for determining the dynamic limits of reservoirs due to their simplicity in analyzing the relationship between reservoir parameters and water inflows. However, these models presume a linear relationship, which may not always be accurate, especially for reservoirs with complex hydrological characteristics. In such cases, more advanced models may be necessary.

To enhance the performance of linear regression models, monthly analysis and filtering methods were applied (as discussed in Sections 3.1 and 4.1.1).

The deviation between actual and forecasted reservoir levels provides valuable information for setting lower and upper bounds. These bounds influence water release and storage policies, ensuring reservoir sustainability and mitigating risks. For instance, if the optimization planning model utilizes the water contained in the equivalent hydro reservoir, the lower bound limits this usage. Conversely, if the model makes decisions regarding storing water, the upper level controls how much water can be stored to meet ecological constraints or strategic behaviors concerning the equivalent reservoir. This method helps implicitly model the uncertainty in the model.

6. Conclusions

This paper introduces analytical techniques for configuring parameters within hydropower optimization models. These techniques include methods based on duration curves and clustering algorithms. Additionally, auxiliary optimization models are utilized to determine water inflows to reservoirs. Moreover, filtering methods based on the Fourier series are applied to linear regression models to establish the limits of equivalent reservoirs (PrU). The simplicity and scalability of these methods make them accessible for practical application.

The benchmark conducted for the 2019 operation in Spain revealed that configuring hydropower parameters through the proposed methods results in a notable alignment between simulated outcomes and the actual operation of the system. This correspondence captures not only the shape of the data but also the specific values of variables such as energy production level and energy storage.

Applying these techniques to a case study in Spain in 2025 yielded promising results. The configured parameters provided flexibility to the optimization model, allowing it to make decisions within reasonable ranges for operation. This demonstrates the potential effectiveness of the approach in addressing the complexities of hydro subsystems in optimization models.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT-4o and ChatGPT-3.5 in order to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Jesús D. Gómez-Pérez: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - Original Draft, Data Curation. **Francisco Labora:** Conceptualization, Formal analysis, Investigation, Writing - Original Draft. **Jesus M. Latorre-Canteli:** Conceptualization, Supervision, Validation, Writing - Review and Editing. **Andres Ramos:** Conceptualization, Supervision, Validation, Writing - Review and Editing, Funding acquisition, Resources.

References

de Amezúa, A.O.C., 2003. Análisis de la gestión hidroeléctrica en el sistema español, in: Anales de mecánica y electricidad, Asociacion de Ingenieros del ICAI. pp. 20–27.

Aradhye, G., Rao, A., Mastan Mohammed, M., 2019. A novel hybrid approach for time series data forecasting using moving average filter and arima-svm, in: Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2018, Volume 2, Springer. pp. 369–381.

Arvanitidis, N.V., Rosing, J., 1970. Optimal operation of multireservoir systems using a composite representation. IEEE Transactions on Power Apparatus and Systems , 327–335.

Arvanitidits, N.V., Rosing, J., 1970. Composite representation of a multireservoir hydroelectric power system. IEEE Transactions on Power Apparatus and Systems , 319–326.

Blom, E., Söder, L., 2022. Comparison of different computational methods and formulations for hydropower equivalents, in: 2022 IEEE 7th International Energy Conference (ENERGYCON), IEEE. pp. 1–6.

Blom, E., Söder, L., 2024. Single-level reduction of the hydropower area equivalent bilevel problem for fast computation. Renewable Energy , 120229.

Blom, E., Söder, L., Risberg, D., 2020. Performance of multi-scenario equivalent hydropower models. Electric power systems research 187, 106486.

CEDEX, . Anuario de aforos 2019 - 2020. modalidades de consulta. URL: https://ceh.cedex.es/anuarioaforos/default.asp.

Cerisola, S., Latorre, J.M., Ramos, A., 2012. Stochastic dual dynamic programming applied to nonconvex hydrothermal models. European Journal of Operational Research 218, 687–697.

ENTSOE, . Entsoe downloads. URL: https://www.entsoe.eu/outlooks/eraa/2023/eraa-downloads/.

ESIOS, . P48 cierre.

ESIOS UF, . Unidades físicas. URL: https://www.esios.ree.es/es/unidades-fisicas. accessed on 2023-12-04.

ESIOS UP, . Unidades programación. URL: https://www.esios.ree.es/es/unidades-de-programacion. accessed on 2023-12-04.

González, P., Villar, J., Díaz, C.A., Campos, F.A., 2013. Hourly energy and reserve joint dispatch with a hydro-thermal technological based representation, in: 2013 10th International conference on the European energy market (EEM), IEEE. pp. 1–8.

Härtel, P., Korpås, M., 2017. Aggregation methods for modelling hydropower and its implications for a highly decarbonised energy system in europe. Energies 10, 1841.

Hassani, H., Mahmoudvand, R., Yarmohammadi, M., 2010. Filtering and denoising in linear regression analysis. Fluctuation and Noise Letters 9, 343–358.

Helseth, A., Mo, B., 2022. Hydropower aggregation by spatial decomposition—an sddp approach. IEEE Transactions on Sustainable Energy 14, 381–392.

Hoffmann, M., Priesmann, J., Nolting, L., Praktiknjo, A., Kotzur, L., Stolten, D., 2021. Typical periods or typical time steps? a multi-model analysis to determine the optimal temporal aggregation for energy system models. Applied Energy 304, 117825.

Iberdrola, 2006. Grandes presas: 110 años del trabajo de Iberdrola. Iberdrola.

Joo, T.W., Kim, S.B., 2015. Time series forecasting based on wavelet filtering. Expert Systems with Applications 42, 3868–3874.

Löschenbrand, M., Korpås, M., 2017. Hydro power reservoir aggregation via genetic algorithms. Energies 10, 2165.

Meng, J., Wang, Y., Guo, H., Ding, Y., 2023. Application of wavelet denoising algorithm in monthly runoff series of fuchun river hydropower station, in: 2023 International Seminar on Computer Science and Engineering Technology (SCSET), IEEE. pp. 719–723.

MITECO, . Boletín hidrológico semanal. URL: https://eportal.miteco.gob.es/BoleHWeb/.

MITECO MAP, . Red hidrográfica básica mdt. URL: https://www.miteco.gob.es/content/dam/miteco/es/cartografia-y-sig/ide/descargas/lyr-Red%20hidrogr%C3%A1fica%20b%C3%A1sica%20MDT%20100x100_tcm30-175726.zip.

MITECO RV, . Sistema de información de redes de seguimiento del estado e información hidrológica y geoportal del ministerio. URL: https://sig.mapama.gob.es/snczi/index.html?herramienta=Presas. accessed on 2023-12-04.

Muhammad, Y.S., Pflug, G.C., 2014. Stochastic vs deterministic programming in water management: the value of flexibility. Annals of Operations Research 223, 309–328.

Niu, S., Insley, M., 2013. On the economics of ramping rate restrictions at hydro power plants: Balancing profitability and environmental costs. Energy Economics 39, 39–52.

Oppenheim, A.V., Willsky, A.S., Nawab, S.H., 1997. Signals and Systems. 2nd ed., Prentice Hall.

de Queiroz, A.R., 2016. Stochastic hydro-thermal scheduling optimization: An overview. Renewable and Sustainable Energy Reviews 62, 382–395.

Ramos, A., Alvarez, E.F., Lumbreras, S., 2022. opentepes: open-source transmission and generation expansion planning. SoftwareX 18, 101070.

Red Eléctrica de España, 2019. Report on the Spanish Electricity System 2019. Technical Report. Red Eléctrica de España. Available at: https://www.sistemaelectrico-ree.es/informe-del-sistema-electrico.

Red Eléctrica de España, 2023. Report on the Spanish Electricity System 2023. Technical Report. Red Eléctrica de España. Available online at https://www.sistemaelectrico-ree.es/informe-del-sistema-electrico.

REE, . Análisis nacional de cobertura del sistema eléctrico peninsular español. URL: https://www.ree.es/sites/default/files/01_ACTIVIDADES/Documentos/informe_os_nov23.pdf.

Ringkjøb, H.K., Haugan, P.M., Solbrekke, I.M., 2018. A review of modelling tools for energy and electricity systems with large shares of variable renewables. Renewable and Sustainable Energy Reviews 96, 440–459.

Stoll, B., Andrade, J., Cohen, S., Brinkman, G., Brancucci Martinez-Anido, C., 2017. Hydropower modeling challenges. Technical Report. National Renewable Energy Lab.(NREL), Golden, CO (United States).

Yue, X., Pye, S., DeCarolis, J., Li, F.G., Rogan, F., Gallachóir, B.Ó., 2018. A review of approaches to uncertainty assessment in energy system optimization models. Energy strategy reviews 21, 204–217.

Zhang, J., Li, H., Chen, D., Xu, B., Mahmud, M.A., 2021. Flexibility assessment of a hybrid power system: Hydroelectric units in balancing the injection of wind power. Renewable Energy 171, 1313–1326.