Electrical Power and Energy Systems 135 (2022) 107464

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

# Piecewise linear approximations for hydropower production function applied on the hydrothermal unit commitment problem

K.V. Santos [a,b,*], E.C. Finardi [b,c]

[a] Federal University of Amazonas, ZIP 69080-900 Manaus, AM, Brazil
[b] Federal University of Santa Catarina/LabPlan, ZIP 88040-900, Florianópolis, SC, Brazil
[c] INESC P&D Brasil, ZIP 11055-300 Santos, SP, Brazil

ARTICLE INFO

*Keywords:*
Hydro Production Function
Piecewise-Linear Models
Hydrothermal Unit Commitment
Mixed-Integer Linear Programming

ABSTRACT

In a centralized-based dispatch market, the network-constrained hydrothermal unit commitment (NCHTUC) determines generation decisions that minimize the expected operating cost, usually in a day-ahead planning horizon. In this context, hydropower offers unique flexibility features, and an adequate representation is crucial for overcoming operation challenges. However, due to computational burden, simplifications on hydropower production function (HPF) modeling are necessary since NCHTUC is a large-scale nonlinear discrete optimization problem. The plant-based HPF piecewise linear approach is the most common simplification presented for real-life cases. In this case, the generating units (GUs) of a plant are aggregated and represented by a single equivalent generator. Although this approach reduces the size and the complexity of the NCHTUC significantly, several operating issues are not considered adequately, especially the forbidden zones and the nonlinearities of the GUs. In this paper, we present a new approach that considers the nonlinearities and forbidden zones of the HPF via aggregation of the GUs and piecewise mixed-integer linear approximation in an innovative way. Experiments are performed in a modified version of the IEEE 118-bus system to verify the approach's effectiveness, analyzed in terms of the relation between computation burden and the HPF approximation quality.

## 1. Introduction

The unit commitment (UC) problem is one of the main activities in the power systems area. This problem aims to find which generators will be active and how much power they will produce to attend a forecasted load over a short-term planning horizon (usually, 24 to 168 hourly time-steps). In systems where hydro and thermal power plants supply several load centers via an extensive electrical network, this problem is commonly described as network-constrained (NC) hydrothermal unit commitment (HTUC). How the NCHTUC is handled depends on the dispatch model adopted in a country or region [1,2]. Decentralized bid-based dispatch models are adopted in wholesale markets, in which the operation is strongly guided by competition between the agents in the purchase and sale of energy.

On the other hand, a centralized cost-based dispatch is usually adopted in less liberalized markets, where the agents do not control their generation. In the centralized dispatch of predominant hydro systems, which is the focus of this paper, one issue that complicates the NCHTUC solution is the hydro production function (HPF) of the generating units (GUs). The HPF, which models the relationship between turbined outflow and electrical power, is a complex nonlinear, nonconvex, and discontinuous function. Another complicating issue is the generation and demand uncertainty. If this aspect is relevant, the resulting NCHTUC should be handled via optimization under uncertainty. As expected, in problems with the predominance of hydro generation, more realistic formulations for the HPF should lead to more complex models. Thus, realistic representations that do not have a sizeable computational impact are interesting for this problem.

Nevertheless, as the focus of this work is related to the HPF, we assess the contributions in a deterministic NCHTUC. An interesting survey about the HTUC (and other UC variants) under uncertainty is detailed in the work [3]. Furthermore, paper [4] presents a strategy based on the Bundle method to deal with the stochastic HTUC, and [5] presents a framework based on Benders decomposition for the NCHTUC. Also, in [6], the robust HTUC problem is modeled by mixed-integer linear programming (MILP).

A significant challenge on the NCHTUC is creating a realistic HPF model, especially in systems with a hydropower predominance. Some of the difficulties are related to the trilinearity (efficiency × head ×

* Corresponding author at: Federal University of Amazonas, ZIP 69080-900 Manaus, AM, Brazil.
*E-mail address:* kennyvinente@ufam.edu.br (K.V. Santos).

https://doi.org/10.1016/j.ijepes.2021.107464
Received 3 March 2021; Received in revised form 7 May 2021; Accepted 28 July 2021
Available online 20 August 2021

**Nomenclature**

*Indices*

| | |
|---|---|
| $j$ | Index of hydro generating units. |
| $t$ | Index of time-steps (hours). |
| $g$ | Index of thermal generating units. |
| $h$ | Index of hydro plants. |
| $b$ | Index of buses. |
| $l$ | Index of transmission lines. |
| $ref$ | Index of reference bus. |
| $i$ | Index of non-forbidden zones. |
| $p$ | Index of polytopes. |

*Variables*

| | |
|---|---|
| $q_{jht}$ | Turbined outflow of the unit $j$, hydro plant $h$ in time-step $t$ ($m^3/s$). |
| $Q_{ht}$ | Turbined outflow of the hydro plant $h$ in time-step $t$ ($m^3/s$). |
| $s_{ht}$ | Spillage of the hydro plant $h$ in time-step $t$ ($m^3/s$). |
| $v_{ht}$ | Volume of the hydro plant $h$ at the beginning in time-step $t$ ($hm^3$). |
| $phu_{jht}$ | Power output of the unit $j$, hydro plant $h$ in time-step $t$ (MW). |
| $php_{ht}$ | Power output of the hydro plant $h$ in time-step $t$ (MW). |
| $\eta_{jht}$ | Turbine efficiency of the unit $j$, hydro plant $h$ in time-step $t$. |
| $nh_{jht}$ | Net head of the unit $j$, hydro plant $h$ in time-step $t$ (m). |
| $fbl_{ht}$ | Forebay level of the hydro plant $h$ in time-step $t$ (m). |
| $trl_{ht}$ | Tailrace level of the hydro plant $h$ in time-step $t$ (m). |
| $hl_{jt}$ | Hydraulic losses of the unit $j$, hydro plant $h$ in time-step $t$ (m). |
| $u_{jht}$ | Binary variable indicating the on/off status of the unit $j$, hydro plant $h$ in time-step $t$. |
| $x_{gt}$ | Binary variable indicating the on/off status of the thermal unit $g$ in time-step $t$. |
| $z_{gt}$ | Binary variable indicating the startup of the thermal unit $g$ in time-step $t$. |
| $w_{gt}$ | Binary variable indicating the shutdown of the thermal unit $g$ in time-step $t$. |
| $pt_{gt}$ | Power output of the thermal unit $g$ in time-step $t$ (p.u. of MW). |
| $pl_{lt}$ | Active power in the transmission line $l$ in time-step $t$ (MW). |
| $\theta_{bt}$ | Voltage angle of the bus $b$ in time-step $t$ (radians). |
| $y_k$ | Binary variable that represents each interval $k$ considered in the PWL model. |
| $\widehat{Q}_k$ | Turbined outflow of the interval $k$ considered in the PWL model ($m^3/s$). |
| $pi_k$ | Power output of the interval $k$ considered in the PWL model (MW). |
| $aphp$ | Power output of the approximated HPF (MW). |
| $aphp^*$ | Value of $aphp$ in the upper limit (MW). |
| $aphp^{**}$ | Value of $aphp$ in the lower limit (MW). |
| $\beta_k$ | Correction factor of the interval $k$ used to estimate the dependence of the volume in the PWL-2 model ($MW/hm^3$). |

*Parameters*

| | |
|---|---|
| **NT** | Number of periods under study (24 h). |
| **NG** | Number of thermal generating units. |
| $\mathbf{C0}_g$ | Unitary variable cost of the thermal unit $g$ ($/MW). |
| $\mathbf{C1}_g$ | Fixed cost of the thermal unit $g$ ($). |
| $\mathbf{C2}_g$ | Startup cost of the thermal unit $g$ ($). |
| $\mathbf{C3}_g$ | Shutdown cost of the thermal unit $g$ ($). |
| $\mathbf{UT}_g$ | Minimum uptime of the thermal unit $g$ (h). |
| $\mathbf{DT}_g$ | Minimum downtime of the thermal unit $g$ (h). |
| $\mathbf{RU}_g$ | Maximum ramp-up of the thermal unit $g$ (MW/h). |
| $\mathbf{RD}_g$ | Maximum ramp-down of the thermal unit $g$ (MW/h). |
| **NH** | Number of hydro plants. |
| $\mathbf{D}_{jh}$ | Constant related to the hydraulic loss function of the unit $j$ and hydro plant $h$ ($s^2/m^5$). |
| $\mathbf{F}_{ih}$ | Constant $i$ related to the forebay function of the hydro $h$. |
| $\mathbf{G}_{ih}$ | Constant $i$ related to the tailrace function of the hydro $h$. |
| $\mathbf{NJ}_h$ | Number of units of the hydro plant $h$. |
| $\mathbf{I}_{ijh}$ | Constant $i$ related to the turbine efficiency function of the unit $j$ and hydro plant $h$. |
| **CQ** | Constant that converts a one hour of water flow ($m^3/s$) to water volume ($hm^3$). |
| $\mathbf{Y}_{ht}$ | Incremental inflow of the hydro plant $h$ in time-step $t$ in ($m^3/s$). |
| $\Omega_h$ | Set of hydro plants that are upstream of the hydro $h$. |
| $\tau_{ih}$ | Water traveling time between hydro plants $i$ and $h$ (h). |
| $\gamma$ | Parameter used to set the initial volume of all hydro plants on the experiments. |
| $\mathbf{U}_b$ | Set of thermal units and hydro plants connected on the bus $b$. |
| **B** | Set of buses of the system. |
| **L** | Set of transmission lines of the system. |
| $\mathbf{L}_b$ | Set of lines connected on the bus $b$. |
| **BM** | Constant that converts MW in p.u. |
| $\mathbf{P}_{bt}$ | Demand for active power on the bus $b$ in time-step $t$ (p.u. of MW). |
| $\mathbf{SR}_t$ | System reserve in time-step $t$ (p.u. of MW). |
| $\mathbf{X}_{ab}$ | Reactance of the line which connects the buses $a$ and $b$ (p.u.). |
| $\mathbf{a}_m$ | Approximated upper bound related to the HPF coefficients for $m = 0,1,2$. |
| $\mathbf{b}_m$ | Approximated lower bound related to the HPF coefficients for $m = 0,1,2$. |
| **N1** | Number of hyperplanes used to represent the upper bound of HPF. |
| **N2** | Number of hyperplanes used to represent the lower bound of HPF. |
| $\mathbf{v}_{ref}$ | Reference volume used in the PWL model ($hm^3$). |
| $\mathbb{I}_h$ | Set of operating zones of HPF of the plant $h$. |
| $\mathbb{P}_h$ | Set of polytopes used in the piecewise linear approximation of HPF of the plant $h$. |
| $\overline{Y}$ | Maximum value of the variable $Y$. |
| $\underline{Y}$ | Minimum value of the variable $Y$. |

turbined outflow) and the forbidden operating zones of the GUs (these zones have more impact on plants with few GUs). Furthermore, systems under centralized cost-based dispatch possess hundreds of hydro GUs and, due to execution time limits, the HPF demands some simplification level. In this sense, according to [7], the plant-based HPF, where the units in a hydro plant are aggregated as an equivalent one, is the most employed modeling strategy for handling the computational effort, even in cases that do not follow a centralized dispatch or NCHTUC [8–13]. The main advantage of using the plant-based concept is that it reduces the optimization problem significantly.

Several works have used the HPF plant-based approach in UC or short-term generation scheduling problems over the years. The paper [8] presents a four-dimensional piecewise linear (PWL) model of HPF as a function of volume ($v$), plant turbined outflow ($Q$), and spillage ($s$), and [13] use this same model for including the effects of river-level and routing constraints on short-term generation scheduling problem. The piecewise linear continuous formulation is obtained through a methodology based on a convex hull (CH). The hyperplanes that are upper

limits on the HPF on $v \times Q$ domain are attained, and post-processing techniques are performed to reduce the average error and to include the spillage. Although the model consists of some nonlinear characteristics, the forbidden zones are not included; therefore, an infeasible operation can be supplied by the approach. Also, the fact of CH is a concave envelope for the HPF, the model can present a null generation even when $Q$ is not zero. Furthermore, approximating a non-convex function by the convex approximation employed in the CH leads to a nonzero linearization errors, no matter how many segments are added. In [9], a plant-based nonlinear approach is employed on the short-term hydro scheduling problem in a deregulated market, where the head-dependency of the HPF is the main subject of study; in [10], an extension of this study to include on–off effects of the hydro plants, avoiding generation at forbidden areas is presented, resulting in a mixed-integer nonlinear programming (MINLP) problem. In turn, [11] uses this same approach for including the start/stop of the units, discharge ramping constraints, and discontinuous operating regions. To accounts for the forbidden zones, these works assume that the HPF has only one discontinuous region. A special issue is that MINLP problems are usually hard to solve, even for small instances. Specifically, in [10] and [11], it is proposed a mixed-integer linear programming (MILP) approach in an attempt to find a starting point for the MINLP or even an optimal solution. The experiments are performed in a reduced configuration of the Portuguese system. Although it is not commented on in these works, the computational complexity of solving an MINLP may be one of the main reasons why the authors did not use a large-size power system.

Another work that uses a plant-based approach is [12], where the nonlinear HPF is linearized using a CH technique, and auxiliary variables model the generation and discharge of the plants. As a result, a space-state model with continuous variables is formulated for the Mid-Columbia hydropower system, consisting of seven hydro plants. At this point, it is also essential to mention the papers [14–16], which develop a plant-based HPF model considering different groups of GUs.

With the recent advances in MILP solvers, the strategy to deal with nonlinear functions is drawing attention. Such strategy is employed in general (but not restricted) to univariate and bivariate functions due to the number of variables and constraints used in each approximation. Although not all nonlinear functions can be rewritten as piecewise linear functions, this reformulation is valid in many problems. In particular [17], presents this methodology in complex processes arising from the gas-network industry. Other references where authors apply MILP for piecewise linear functions are [18]. In addition, the works [19–23] apply this methodology to formulate the HPF in various hydro generation problems. Recently, optimization solvers are including such an approach, which allows handling certain types of nonlinear optimization problems, including GUROBI [24].

Recent works use the advances of the PWL formulations and MILP approaches on the HTUC problem. For example, [25] inserts the HPF nonlinearities dynamically by PWL functions, including the effects of penstock and tailrace losses on the net head. Then, the problem is solved by a two-stage model, where the first stage is a MILP and the second stage is a linear programming (LP) problem. In [26], the authors present MILP formulations to deal with the HPF on the short-term scheduling problem based on PWL approximations in a small-size three-plant system. Although these formulations possess high precision, the computational effort for a large-scale instance is a challenge. Therefore, an intermediate approach with an acceptable level of accuracy for the HPF with a slight increase in the computational effort is attractive.

This paper proposes different plant-based formulations for the HPF, including nonlinearities and discontinuities related to the forbidden operating zones through mixed-integer formulations, aiming to find a formulation with reasonable HPF accuracy and a slight increase in the computational effort. In other words, our main contribution is to present an intermediate approach (something between the individual representation and the plant-based approach), which includes details present in the individual representation and the reduced computational effort characteristic of the plant-based approach. Initially, we present a two-dimensional mixed-integer formulation based on the CH model for each non-forbidden zone. Then, from a detailed analysis of the HPF, correlating the state variables' dependence and the planning horizon under consideration, one and bidimensional piecewise linear models are formulated. Finally, we employ a similar approach as [6], in which the experiments are performed in a modified version of the IEEE-118 bus system with 15 hydro plants. Each HPF proposed is evaluated and compared with the classic CH approach in terms of computational burden and nonlinear HPF approximation quality.

The rest of the paper is organized as follows: the next section presents the HPF modeling proposed in this work; Section 4 contains the mathematical formulation of the centralized cost-based dispatch NCHTUC; Section 5 presents the experiments performed, and conclusions are depicted in Section 6.

## 2. Hydropower production function modeling

The power output of a GU is a nonlinear function that depends on the turbined outflow, net head, and the (turbine-generator) efficiency as follows:

$$phu_{jht} = 9.81 \times 10^{-3} \eta_{jht} nh_{jht} q_{jht}. \tag{1}$$

The net head is represented as follows:

$$nh_{jht} = fbl_{ht} - trl_{ht} - hl_{jht}. \tag{2}$$

The following function represents the hydraulic head loss:

$$hl_{jht} = \mathbf{D}_{jh} q_{jht}^2. \tag{3}$$

The constant $\mathbf{D}_{jh}$ depends on the physical characteristics of the penstock and thrash hacks. In this work, the forebay level is given by the following fourth-degree polynomial:

$$fbl_{ht} = \mathbf{F}_{0h} + \mathbf{F}_{1h} v_{ht} + \mathbf{F}_{2h} v_{ht}^2 + \mathbf{F}_{3h} v_{ht}^3 + \mathbf{F}_{4h} v_{ht}^4. \tag{4}$$

The tailrace level is also modeled as a fourth-degree polynomial:

$$trl_{ht} = \mathbf{G}_{0h} + \mathbf{G}_{1h}(Q_{ht} + s_{ht}) + \mathbf{G}_{2h}(Q_{ht} + s_{ht})^2 + \mathbf{G}_{3h}(Q_{ht} + s_{ht})^3 + \mathbf{G}_{4h}(Q_{ht} + s_{ht})^4. \tag{5}$$

The GU efficiency is a nonlinear function expressed in the Hill diagram [27]. In this work, we use the following polynomial approximation model:

$$\eta_{jht} = \mathbf{I}_{0jh} + \mathbf{I}_{1jh} q_{jht} + \mathbf{I}_{2jh} nh_{jht} + \mathbf{I}_3 q_{jht} nh_{jht} + \mathbf{I}_4 q_{jht}^2 + \mathbf{I}_5 nh_{jht}^2. \tag{6}$$

The power generation of the hydro plant is the sum of the power output of all GUs:

$$php_{ht} = \sum_{j=1}^{\mathbf{NJ}_h} phu_{jht}. \tag{7}$$

Similarly, the turbined outflow of the hydro plant is the sum of the individual turbined outflow of all GUs:

$$Q_{ht} = \sum_{j=1}^{\mathbf{NJ}_h} q_{jht}. \tag{8}$$

As commented previously, a common approach to represent the HPF is to use a plant-based model, which models the plant via a single equivalent GU. In this case, the equivalent unit possesses maximum power and turbined outflow given by the sum of the individual GU capacity. However, even where only identical GUs exist, it is impossible to obtain a precise analytic expression for the HPF in the plant-based approach since the resulting $php_{ht}$ depends on the generation policy (i.e., 0–1 status and the respective power dispatch) chosen for each unit. Therefore, the power generation in the operating point $(\mathbf{Q}_{ht}^k, \mathbf{v}_{ht}^k)$ can only be obtained by solving an auxiliary problem, typically an optimization

model. For instance, considering null spillage, the output power of the plant, $php_{ht}$, operating with fixed values of $\mathbf{Q}_{ht}^{k}$ and $\mathbf{v}_{ht}^{k}$, is found by solving the following MINLP problem:

$$php_{ht}\left(\mathbf{Q}_{ht}^{k},\mathbf{v}_{ht}^{k}\right)=max\sum_{j=1}^{\mathbf{NJ}_h}phu_{jht} \tag{9}$$

$$\text{s.t.}:\sum_{j=1}^{\mathbf{NJ}_h}q_{jht}=\mathbf{Q}_{ht}^{k} \tag{10}$$

$$nh_{jht}=\sum_{p=0}^{4}\mathbf{F}_{ph}\cdot\left(\mathbf{v}_{ht}^{k}\right)^{p}-\sum_{p=0}^{4}\mathbf{G}_{ph}\cdot\left(\mathbf{Q}_{ht}^{k}\right)^{p}-\mathbf{D}_{jh}q_{jht}^{2} \tag{11}$$

$$\eta_{jht}=\mathbf{I}_{0jh}+\mathbf{I}_{1jh}q_{jht}+\mathbf{I}_{2jh}nh_{jht}+\mathbf{I}_{3jh}q_{jht}nh_{jht}+\mathbf{I}_{4jh}q_{jht}^{2}+\mathbf{I}_{5jh}nh_{jht}^{2} \tag{12}$$

$$phu_{jht}=9.81\times10^{-3}\eta_{jht}nh_{jht}q_{jht} \tag{13}$$

$$u_{jht}\cdot\underline{\mathbf{QU}}_{ht}\leqslant q_{jht}\leqslant u_{jht}\cdot\overline{\mathbf{QU}}_{ht} \tag{14}$$

$$u_{jht}\in\{0,1\} \tag{15}$$

$$j=\{1,2,\ldots,\mathbf{NJ}_h\} \tag{16}$$

As an illustrative example, the HPF of a 3-unit plant is presented in Fig. 1, wherein (a) the impact of the forbidden zones is not considered. Each generation value in this figure is obtained by solving the previous MINLP optimization problem.

As can be seen, there are significant differences between HPFs due to the consideration of operating zones. Therefore, in the rest of this paper, all proposed HPFs are formulated considering the plant-based model and peculiar modeling issues. For example, hydro plants with large reservoirs, where variations in the volume are irrelevant in the short-term planning, are modeled by a one-dimensional HPF. Also, we will omit the subscripts $h$ and $t$ to simplify the notation.

### 2.1. Envelope models

For reducing the computational effort associated with the nonlinear individual HPF, [8] proposes a piecewise linear model that considers the effects of volume, turbined outflow, and spillage. The main advantage of this technique is that the resulting model is linear, where one of the disadvantages is that it does not consider the forbidden zones. Experiments for this case also present other numerical aspects; for example, many of the hyperplanes obtained by the Convex Hull algorithm are similar numerically, and the user does not control the number of hyperplanes obtained. Thus, in this work, we propose to find the hyperplanes through an optimization problem, which is described below. First, consider a set $S$ of points in the grid and a subset $S_1$[1], we solve the following optimization problem for each element $m\in S_1$:

$$min\sum_{k\in S}\left[php\left(\mathbf{Q}^{k},\mathbf{v}^{k}\right)-\left(a_0\mathbf{Q}^{k}+a_1\mathbf{v}^{k}+a_2\right)\right]^{2} \tag{17}$$

$$\text{s.t.:}\ \left(a_0\mathbf{Q}^{k}+a_1\mathbf{v}^{k}+a_2\right)-php\left(\mathbf{Q}^{k},\mathbf{v}^{k}\right)\geqslant0,\forall k\in S \tag{18}$$

$$\left(a_0\mathbf{Q}^{m}+a_1\mathbf{v}^{m}+a_2\right)-php(\mathbf{Q}^{m},\mathbf{v}^{m})=0 \tag{19}$$

Note that the spillage is not considered in this model, although it can include a post-processing strategy [8]. The optimization problem can be infeasible since HPF is nonconcave. However, if the problem is feasible, the resulting hyperplane is an upper limit for $php$. Considering the interest in finding adequate representations for each permutation of extreme points of $(Q,v)$[2], the minimum number of elements in $S_1$ is 4. The resulting HPF model, called Convex Hull (CH) model, is formulated as follows:

$$aphp\leqslant\mathbf{a}_{0k}Q+\mathbf{a}_{1k}v+\mathbf{a}_{2k},k=1,\ldots,\mathbf{N1} \tag{20}$$

$$0\leqslant Q\leqslant\overline{q}\cdot\mathbf{NJ}, \tag{21}$$

where $\mathbf{N1}$ is the number of hyperplanes obtained after solving the optimization problem described by (17)-(19) for all m $\in S_1$, which is less than or equal to the cardinality of $S_1$. One disadvantage of this formulation is that this model is a good approximation only if the nonlinear HPF is concave. Moreover, this formulation does not consider the forbidden zones and, as Equation (20) is an upper bound, $aphp$ can be zero even for a nonzero $Q$. An alternative to correct the null generation when $Q>0$ based on [11] is to include a binary variable $u$ that represents the on/off status of the hydro plant and adding the following pair of constraints:

$$\underline{php}\cdot u\leqslant aphp\leqslant\overline{php}\cdot u, \tag{22}$$

$$\underline{q}\cdot u\leqslant Q\leqslant\overline{q}\cdot\mathbf{NJ}. \tag{23}$$

The resulting model for the HPF, obtained by combining Eqs. (20), (22)–(23) is called the UCH model. Note that formulation does not consider the forbidden zones. To do that, we propose a mixed-integer bilinear formulation for the HPF, called Piecewise Convex Hull (PCH) model. The upper limits are evaluated for each non-forbidden zone in the same way as performed for the CH model, and lower bounds on HPF are imposed. Defining $\mathbb{I}$ as the set of non-forbidden zones, where $i\in\mathbb{I}$ has the form $i=[Q_{i1},Q_{i2}]$, $Q_{i1}<Q_{i2}$ and represents a non-forbidden zone on $Q$, the PCH model is formulated as follows:

$$aphp\leq\sum_{i\in\mathbb{I}}\mathbf{a}_{0ki}\widehat{Q}_i+\mathbf{a}_{1ki}v\cdot y_i+\mathbf{a}_{2ki}y_i,k=1,\ldots,\mathbf{N1}_i \tag{24}$$

$$aphp\geq\sum_{i\in\mathbb{I}}\mathbf{b}_{0ki}\widehat{Q}_i+\mathbf{b}_{1ki}v\cdot y_i+\mathbf{b}_{2ki}y_i,k=1,\ldots,\mathbf{N2}_i \tag{25}$$

$$Q=\sum_{i\in\mathbb{I}}\widehat{Q}_i \tag{26}$$

$$Q_{i1}\cdot y_i\leq\widehat{Q}_i\leq Q_{i2}\cdot y_i,\forall i\in\mathbb{I} \tag{27}$$

$$\sum_{i\in\mathbb{I}}y_i\leq1 \tag{28}$$

$$y_i\in\{0,1\},\forall i\in\mathbb{I}, \tag{29}$$

where $\mathbf{N1}_i$, $\mathbf{N2}_i$ are the number of hyperplanes used to limit $aphp$ from above and below, respectively, $Q_{i1}$ and $Q_{i2}$ are the extreme points in $Q$ for each non-forbidden zone, and $y$ is a binary variable that represents each non-forbidden zone. For example, in the case of Fig. 1(b), which presents three non-forbidden zones, the coefficients of hyperplanes of (25) are obtained through the following optimization problem:

$$min\sum_{k\in S}\left[php\left(\mathbf{Q}^{k},\mathbf{v}^{k}\right)-\left(b_0\mathbf{Q}^{k}+b_1\mathbf{v}^{k}+b_2\right)\right]^{2} \tag{30}$$

s.t.:

$$php\left(\mathbf{Q}^{k},\mathbf{v}^{k}\right)-\left(b_0\mathbf{Q}^{k}+b_1\mathbf{v}^{k}+b_2\right)\geqslant0,\forall k\in S \tag{31}$$

$$php(\mathbf{Q}^{m},\mathbf{v}^{m})-(b_0\mathbf{Q}^{m}+b_1\mathbf{v}^{m}+b_2)=0, \tag{32}$$

where $S$, $S_1$, and $m\in S_1$ are the same used in Eqs. (17)–(19).

[1] The subset $S_1$ contains points of interest from the $v\times Q$ grid, such that the value of linear HPF for each element of this subset is the same as the nonlinear HPF.

[2] As example, is expected a representation which has a minimal (if possible, null) error for $\overline{Q}$, and 0 for a null $Q$.

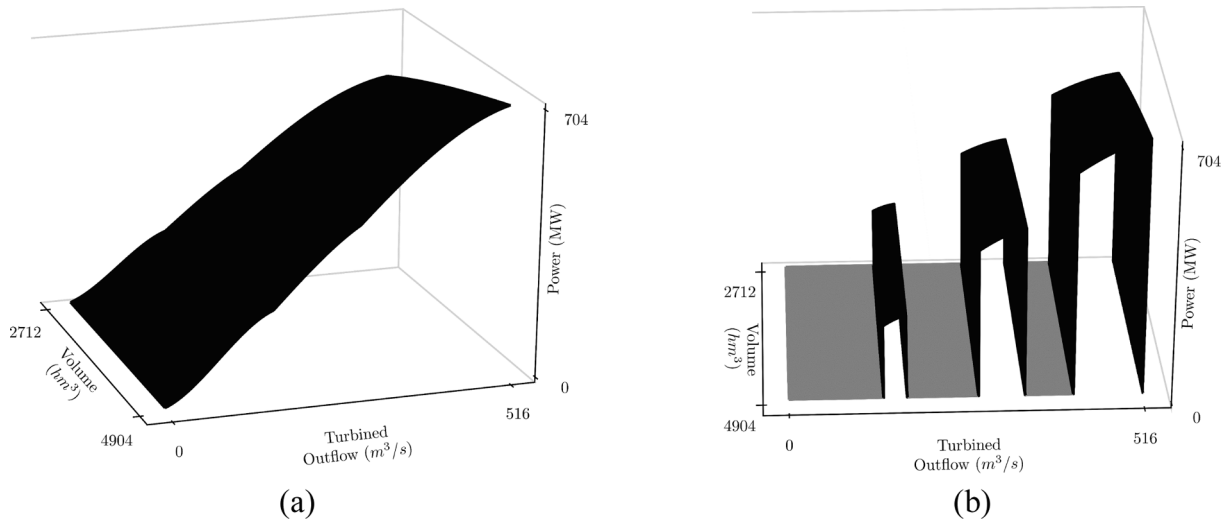



**Fig. 1.** HPF of a 3-unit plant, where (a) ignores the forbidden zones and (b) accounts for these zones.

### 2.2. Mixed-Integer linear approximation models

Another approach to represent nonlinear functions is by mixed-integer piecewise linear models [28]. This section presents a mixed-integer formulation that approximates the nonlinear HPF by piecewise-linear equations on each non-forbidden zone. The focus is on two types: a one-dimensional model as a function of $Q$ and the two-dimensional models as a function of $Q$ and $v$. The one-dimensional HPF is appropriated if the volume variations are sufficiently slight in the planning horizon (it is valid for plants with large reservoirs). For example, Fig. 2(a) presents the nonlinear HPF of Fig. 1 for a fixed volume $\mathbf{v}_0$, equal to 60% of the useful volume, i.e., $\mathbf{v}_0 = 0.6\left(\overline{v} - \underline{v}\right) + \underline{v}$. On the other hand, Fig. 2(b) details the HPF in the turbined outflow range [415.5–516] $m^3/s$ for three volumes, $\mathbf{v}_0, \mathbf{v}_1 = \mathbf{v}_0 - \mathbf{CQ}\cdot\mathbf{NT}\cdot\overline{Q}$ and $\mathbf{v}_2 = \mathbf{v}_0 + \mathbf{CQ}\cdot\mathbf{NT}\cdot\mathbf{Y}$.[3]

In Fig. 2(b), the maximum difference between the HPFs is less than 1.6%, which is the maximum expected error for a model that ignores the volume.

For representing the HPF through a mixed-integer piecewise linear model as [23–24], it is first necessary to define the polytopes used on the linear approximations. Then, let be $\mathbb{P}$ the set of polytopes used in the mixed-integer piecewise linear approximation of HPF and $p \in \mathbb{P}$ on the form $p = [Q_{p1}, Q_{p2}]$, $Q_{p1} < Q_{p2}$. Imposing that this formulation must represent the HPF in all non-forbidden zones, each polytope is contained in one of the non-forbidden zones. Thus, the one-dimensional piecewise linear model (PWL-1) for HPF considering a reference volume $\mathbf{v}_{ref}$ is given by:

$$Q = \sum_{p\in\mathbb{P}} \hat{Q}_p \tag{33}$$

$$aphp = \sum_{p\in\mathbb{P}} pi_p \tag{34}$$

$$pi_p = php(Q_{p1}, \mathbf{v}_{ref})\cdot y_p + \frac{php(Q_{p2}, \mathbf{v}_{ref}) - php(Q_{p1}, \mathbf{v}_{ref})}{Q_{p2} - Q_{p1}}(\hat{Q}_p - Q_{p1}\cdot y_p), \forall p \in \mathbb{P} \tag{36}$$

$$Q_{p1}\cdot y_p \le \hat{Q}_p \le Q_{p2}\cdot y_p, \forall p \in \mathbb{P} \tag{35}$$

$$\sum_{p\in\mathbb{P}} y_p \le 1 \tag{37}$$

$$y_p \in \{0,1\}, \forall p \in \mathbb{P}, \tag{38}$$

where $\mathbf{v}_{ref}$ is a value between $\mathbf{v}_1$ and $\mathbf{v}_2$ (usually, the mean value is chosen). Note that the number of intervals is related to the quality of the linear approximation. Thus, more intervals lead to better representations at the price of an increase in the computational effort.

The second formulation for HPF considers the effects of volume. First, it can represent the HPF from Fig. 2 as a piecewise bidimensional model using formulations given in [28]. However, the high number of binary variables needed to represent the functions and the approximation precision does not make the approach attractive. So instead, we are proposing an alternative formulation that uses a lower number of binary variables.

Analyzing Fig. 2(b), the three HPFs have similar graphics, differentiating by a factor $\beta_p(Q)\cdot v$. Considering $\beta_p(Q)$ as a constant, we can derive a two-dimensional piecewise formulation for the HPF (PWL-2). In this work, the equation used to define beta is:

$$\beta_p(Q) = \frac{php(Q_{p1}, \overline{v}) - php(Q_{p1}, \mathbf{v}_{ref})}{\overline{v} - \mathbf{v}_{ref}}, \tag{39}$$

[3] The planning horizon is one day discretized in hours, i.e., $\mathbf{NT} = 24$, $\mathbf{CQ} = 3600\cdot10^{-6}$, $\mathbf{Y} = 1000$ $m^3/s$ and $\overline{Q} = 516$ $m^3/s$.

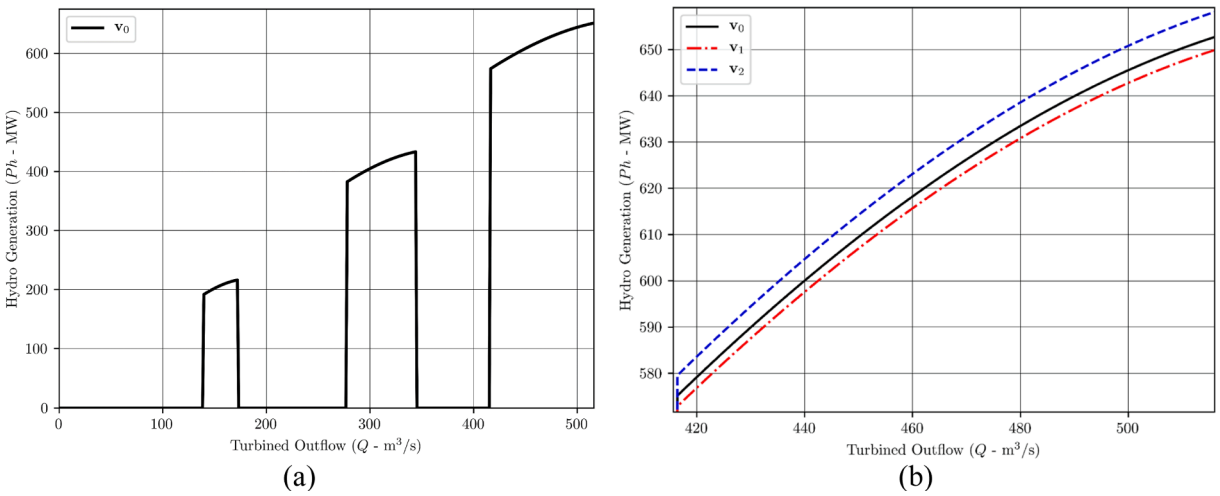



**Fig. 2.** (a) HPF considering fixed volume equal to $v_0$, and (b) detail of HPF in the turbined outflow range [415.5, 516] $m^3/s$ for three fixed volumes.

and the PWL-2 model is formulated by Eqs. (33)–(38), where equation (36) is replaced by:

$$pi_p = php(Q_{p1}, \mathbf{v}_{ref})y_p + \frac{php(Q_{p2}, \mathbf{v}_{ref}) - php(Q_{p1}, \mathbf{v}_{ref})}{Q_{p2} - Q_{p1}}(\hat{Q}_p - Q_{p1} \cdot y_p) + \beta_p(v - \mathbf{v}_{ref})y_p, \forall p \in \mathbb{P}. \tag{40}$$

Finally, we emphasize that all products between continuous and binary variables can be reformulated as mixed-integer linear constraints, resulting in mixed-integer linear models for all proposed HPF approximations.

### 2.3. Numerical example

Considering the HPF formulations presented, we show a numerical example to familiarize the reader with the formulations and main characteristics. The following data are used: $\mathbf{v}_2 = 4{,}113.7$ $hm^3$, $\mathbf{v}_1 = 3{,}982.8$ $hm^3$, $v_{ref} = 0.6 \times (\mathbf{v}_2 - \mathbf{v}_1) + \mathbf{v}_1$. The validation points are obtained through the cartesian product between 200 equally spaced points in $Q$, in which $i \in \mathbb{I}$, with $\mathbb{I} = [(138.5, 172), (277, 344), (415.5, 516)]$, and ten equally spaced points in $v$ on the range $[\mathbf{v}_1, \mathbf{v}_2]$. The approximation errors are calculated based on the following definitions.

**Definition 3.1.** (*i*) *The relative error* ($R_{err}$) *at a point* $x_0$ *is the absolute error divided by the absolute value of the function at* $x_0$, *i.e.*,

$$R_{err}(x_0) = \left|\frac{f(x_0) - \tilde{f}(x_0)}{f(x_0)}\right|, f(x_0) \neq 0. \tag{41}$$

(ii) Given an interval $[x_0, x_1]$, the normalized accumulated error ($NA_{err}$) is the sum of errors at all valid points of the function $f(x)$ in this range divided by the number of interval elements.

To evaluate the normalized accumulated error is necessary to consider each type of approximation used on HPF. For example, in the envelope model, we use two types of error: the first, which measure the difference between the value of the approximation on the upper bound and the value of the nonlinear HPF, called *ub_error*; the second error, related to the difference between the nonlinear HPF and the approximation on the lower limit, called *lb_error*.

Considering a set $\mathbb{Q}$ with $N$ points in interval $[Q_1, Q_2]$, the *ub_error* is obtained as:

$$\frac{1}{N}\sum_{k \in \mathbb{Q}} \max\left\{\left|\frac{php(k, \mathbf{v}) - aphp*(k, \mathbf{v})}{php(k, \mathbf{v})}\right|, 0\right\}, \tag{42}$$

where $k$ is an element of the set $\mathbb{Q}$ and $\mathbf{v}$ is a fixed volume.

The *lb_error*, which is associated with the PCH model, is given by:

$$\frac{1}{N}\sum_{k \in \mathbb{Q}} \max\left\{\left|\frac{php(k, \mathbf{v}) - aphp**(k, \mathbf{v})}{php(k, \mathbf{v})}\right|, 0\right\}. \tag{43}$$

In this equation, *aphp** is the value of *aphp* in the upper limit of (20), and *aphp*** is the value of *aphp* in the lower limit of (24).

Note that we do not measure the *lb_error* for the CH model once this model does not have a lower limit (the lower bound is equal to 0). Finally, for the piecewise models, considering the interval $[Q_1, Q_2]$ divided into $\mathbf{P}$ segments, where $\mathbf{P}_p$ is the subinterval related to segment $p$ and each element $k$ belonging to one of the segments, an expression for the $NA_{err}$ is given by:

$$\frac{1}{N}\sum_{p=0}^{\mathbf{P}-1}\sum_{k \in \mathbb{P}_p}\left|\frac{php(k, \mathbf{v}) - pi_p(k, \mathbf{v})}{php(k, \mathbf{v})}\right|. \tag{44}$$

Now, we can assess each HPF approximation presented in this section, using the following parameters for each model:

- CH: elements of $S_1$ set as the result of the cartesian product between [0, 172, 344, 516] on $Q$ and $[\mathbf{v}_1, (\mathbf{v}_1 + \mathbf{v}_2)/2, \mathbf{v}_2]$ on $v$. Type of error: *ub_error*;
- PCH: $\mathbb{I} = [(138.5, 172), (277, 344), (415.5, 516)]$. Each subset $S_{1i}$ contain elements obtained through the cartesian product between $i \in \mathbb{I}$ and $[\mathbf{v}_1, (\mathbf{v}_1 + \mathbf{v}_2)/2, \mathbf{v}_2]$. Type of error: *ub_error* and *lb_error*;
- PWL: $\mathbb{P} = [(138.5, 172), (277, 311), (311,344), (415.5, 466), (466, 516)]$. Type of error: $NA_{err}$ from equation (44).

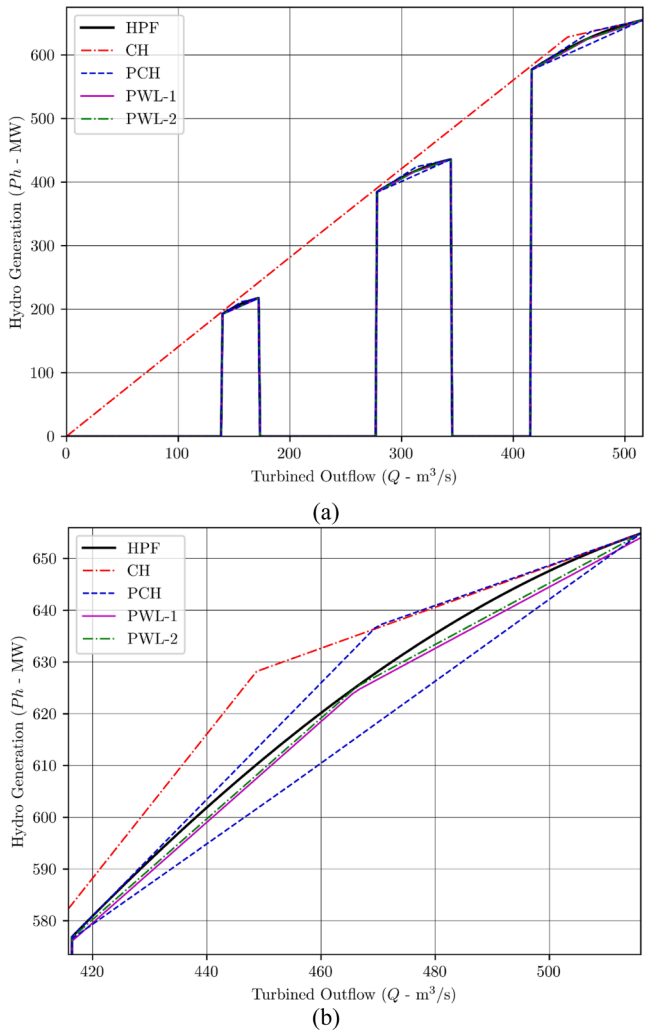



**Fig. 3.** (a) Different HPF for 3-unit plant considering fixed volume, and (b) detail on the turbined outflow range [415.5, 516] $m^3/s$.

**Table 1**
HPF approximations and their characteristics.

| Type | Neq | Nbvar | Npcv | Error in non-forbidden zone # | | |
|---|---|---|---|---|---|---|
| | | | | 1 | 2 | 3 |
| CH | 6 | 0 | 0 | 5.60% | 5.19% | 1.27% |
| PCH | 30 | 3 | 30 | 0.46–1.01% | 0.46–0.25% | 0.46–0.25% |
| PWL-1 | 5 | 5 | 0 | 1.04% | 1.03% | 1.03% |
| PWL-2 | 5 | 5 | 5 | 1.04% | 0.39% | 0.39% |

The graphics of the HPFs of this example are shown in Fig. 3, and the results are summarized in Table 1, where Neq is the number of planes used to represent *aphp*, Nbvar is the number of binary variables used in the model, and Npcv is the number of products between a continuous and a binary variable of the model.

From these results, we observe that PCH and PWL models are superior for representing the HPF.

## 3. Mathematical formulation of network-constrained hydrothermal unit commitment

The NTHTUC considered in this study is related to a centralized tight-pool-based model. The thermal operating generation cost composes the objective function. The planning horizon is composed of 24 time-steps of one hour, and the transmission system is modeled as a classic DC model.

Also, considering that we have three different types of formulations for the HPF, to simplify the notation, we will describe the HPF in this optimization problem as a function of the state variables, where details of this formulation will be presented throughout this section. The formulation of this problem is given by:

$$min \sum_{t=1}^{\mathbf{NT}} \sum_{g=1}^{\mathbf{NG}} \left(\mathbf{C0}_g pt_{gt} + \mathbf{C1}_g x_{gt} + \mathbf{C2}_g z_{gt} + \mathbf{C3}_g w_{gt}\right) \tag{45}$$

$$s.t.: \sum_{g \in \mathbf{U}_b} \mathbf{BM} \cdot pt_{gt} + \sum_{h \in \mathbf{U}_b} \mathbf{BM} \cdot aphp_{ht} - \sum_{l \in \mathbf{L}_b} pl_{lt} = \mathbf{P}_{bt}, \forall b \in \mathbf{B}, t \in \{1, \ldots, \mathbf{NT}\} \tag{46}$$

$$pl_{lt} = (\theta_{at} - \theta_{bt})/\mathbf{X}_{ab}, \forall l = (a,b) \in \mathbf{L}, t \in \{1, \ldots, \mathbf{NT}\} \tag{47}$$

$$\underline{pl}_l \leqslant pl_{lt} \leqslant \overline{pl}_l, \forall \in \mathbf{L}, t \in \{1, \ldots, \mathbf{NT}\} \tag{48}$$

$$\underline{\theta}_b \leqslant \theta_{bt} \leqslant \overline{\theta}_b, \forall b \in \mathbf{B}, t \in \{1, \ldots, \mathbf{NT}\} \tag{49}$$

$$\theta_{ref,t} = 0, t \in \{1, \ldots, \mathbf{NT}\} \tag{50}$$

$$\mathbf{BM} \cdot \left( \sum_{g=1}^{\mathbf{NG}} \left( \overline{pt}_g - pt_{gt} \right) x_{gt} + \sum_{h=1}^{\mathbf{NH}} \left( \overline{php}_h - aphp_{ht} \right) \right) \geqslant \mathbf{SR}_t, t \in \{1, \ldots, \mathbf{NT}\} \tag{51}$$

$$z_{gt} - w_{gt} = x_{gt} - x_{g,t-1}, g \in \{1, \ldots, \mathbf{NG}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{52}$$

$$\sum_{k=t-\mathbf{UT}_g+1}^{t} z_{gk} \leqslant x_{gk}, g \in \{1, \ldots, \mathbf{NG}\}, t \in \{\mathbf{UT}_g, \cdots, \mathbf{NT}\} \tag{53}$$

$$\sum_{k=t-\mathbf{DT}_g+1}^{t} w_{gk} \leqslant 1 - x_{gk}, g \in \{1, \ldots, \mathbf{NG}\}, t \in \{\mathbf{DT}_g, \cdots, \mathbf{NT}\} \tag{54}$$

$$pt_{gt} - pt_{g,t-1} \leqslant \mathbf{RU}_g x_{g,t-1} + \underline{pt}_g z_{gt}, g \in \{1, \ldots, \mathbf{NG}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{55}$$

$$pt_{g,t-1} - pt_{gt} \leqslant \mathbf{RD}_g x_{gt} + \underline{pt}_g w_{gt}, g \in \{1, \ldots, \mathbf{NG}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{56}$$

$$\underline{pt}_g x_{gt} \leqslant pt_{gt} \leqslant \overline{pt}_g x_{gt}, g \in \{1, \ldots, \mathbf{NG}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{57}$$

$$aphp_{ht} = php(v_{ht}, Q_{ht}, s_{ht}), h \in \{1, \ldots, \mathbf{NH}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{58}$$

$$v_{ht} + \mathbf{CQ}\left( Q_{ht} + s_{ht} - \sum_{i \in \Omega_h} \left( Q_{i,t-\tau_{ih}} + s_{i,t-\tau_{ih}} \right) \right) = v_{h,t-1} + \mathbf{CQ} \cdot \mathbf{Y}_{ht}, h \in \{1, \ldots, \mathbf{NH}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{59}$$

$$v_{h,\mathbf{NT}} \geqslant \mathbf{v}_h^*, h \in \{1, \ldots, \mathbf{NH}\} \tag{60}$$

$$\underline{v}_h \leqslant v_{ht} \leqslant \overline{v}_h, h \in \{1, \ldots, \mathbf{NH}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{61}$$

$$Q_{ht} \in \left[ \underline{Q}_h, \overline{Q}_h \right], h \in \{1, \ldots, \mathbf{NH}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{62}$$

$$0 \leqslant s_{ht} \leqslant \overline{s}_h, h \in \{1, \ldots, \mathbf{NH}\}, t \in \{1, \ldots, \mathbf{NT}\} \tag{63}$$

$$u_{ht}, z_{gt}, w_{gt}, x_{gt} \in \{0,1\}, g \in \{1, \ldots, \mathbf{NG}\}, h \in \{1, \ldots, \mathbf{NH}\}, t \in \{1, \ldots, \mathbf{NT}\}. \tag{64}$$

The objective function is composed of the thermal costs. For the hydro plants, we are not considering the operational costs since one of the main objectives of this work is to evaluate the different HPF modeling types. Nevertheless, centralized tight-pool-based models commonly employ the piecewise convex linear functions to estimate future water value obtained through long-term operational planning models [30].

Eqs. (46)–(50) are related to the DC network model, where the transmission line $l$ connects the buses $a$ and $b$ has reactance $\mathbf{X}_{ab}$. Eq. (51) is a system reserve constraint. The thermal modeling is represented by Eqs. (52)–(57), based on the model presented in [31]. Eq. (58) represents the aggregated HPF model, which can be one of the models of Section 3. Eq. (59) is the water balance equation, and Eq. (60) imposes volume targets at the end of the planning horizon. Finally, Eqs. (61)–(64) are limits on variables, where Eq. (62) depends on the type of HPF model used.

## 4. Case study

In this paper, the different formulations for the HPF on the NCHTUC problem are assessed in a modified version of the IEEE 118-bus system [32], where 15 hydro plants replace a subset of 14 thermal plants. The optimization model has been developed and implemented in Python and solved using the optimization solver GUROBI [24], performed on a computer with a Ryzen 9 3900X 12-core processor with 16 GB of RAM. Each simulation has a 10-minute time limit, and an optimality gap of 1% is imposed. The modifications and hydro data can be found in *https://github.com/kennyvinente/scucdata*, and the following experiments are realized:

- Case 1: CH formulation for all hydro plants, i.e., Eqs. (58) and (62), will be replaced by (20) and (21), respectively.
- Case 2: UCH formulation for all hydro plants, i.e., Eq. (58) will be replaced by Eqs. (20)–(22), and equation (62) will be replaced by (23).
- Case 3: PCH formulation for all hydro plants, i.e., Eqs. (58) and (62) will be replaced by Eqs. (24)–(29).
- Case 4: PWL formulation for all hydro plants, i.e., Eqs. (58) and (62) will be replaced by Eqs. (33)–(38) for PWL-1 model and Eqs. (33)–(35),(37)–(40) for PWL-2 model. The choice between PWL-1 and PWL-2 formulation for each plant will be presented in this section.

All the experiments are performed considering three initial volumes on the reservoirs (30, 60, and 90% of volume) and two inflows ($\mathbf{Y}_0$ and $\mathbf{Y}_1$), totalizing 24 different experiments. Regarding the use of binary variables in FPH, Case 1 do not use binary variables; Case 2 includes one binary variable for each hydro plant; Case 3 includes one binary variable for each non-forbidden zone; and Case 4 is a model in which it is possible to set the precision of the HPF, at the price of using a large number (possibly more than used in Case 3) of binary variables. To choose the number of polytopes in Case 4, we developed a methodology that evaluates each hydro capacity concerning the hydro system; a minimal error is imposed depending on this participation. Also, to determine the partitions of the polytopes for each hydro, a variation of the Douglas-Pecker algorithm [33] is used. These details are described in the next section.

### 4.1. Generation of the set of polytopes

As mentioned before, the precision of HPF on PWL formulations depends on the set of polytopes $\mathbb{P}$. Therefore, a pre-processing step to generate this set is performed considering the specific characteristics of the hydro plants. Specifically, the dependency of volume and the total participation on the installed hydro generation capacity are analyzed, and after, the polytopes are generated. For the system used in this paper, Table 2 shows this information for the hydro plants, where the hydro generation capacity is 42.65% of the system.

In Table 2, PHH is the participation of the hydro plant concerning the capacity of all hydro plants, PHP is the participation of the hydro plant to the capacity of the system, and $\Delta\text{php}(Q)$ is a factor that measures the influence of volume on the HPF, given by the following equation:

$$\Delta\text{php}(\mathrm{Q}) = \frac{\mathbf{PH}(\mathrm{Q}, \mathbf{v}_2) - \mathbf{PH}(\mathrm{Q}, \mathbf{v}_1)}{\mathbf{PH}(\mathrm{Q}, \mathbf{v}_0)} \tag{65}$$

where $\mathbf{v}_0$ is the initial volume used in the experiment, $\mathbf{v}_1 = \max(\mathbf{v}_{\min}, \mathbf{v}_0 - \mathbf{NT} \times \mathbf{CQ} \times \mathbf{Q}_{\max})$ is the minimum volume available of the reservoir, and $\mathbf{v}_2 = \min(\mathbf{v}_{\max}, \mathbf{v}_0 + \mathbf{NT} \times \mathbf{CQ} \times (\mathbf{Y} + \mathbf{Q}_{\text{ups}}))$ is the maximum volume available of the reservoir, $\mathbf{Q}_{\text{ups}}$ is the maximum outflow of the reservoirs on the upstream and $\gamma$ is a parameter used to set the initial volume, i.e., $\mathbf{v}_0 = \gamma \times (\mathbf{v}_{\max} - \mathbf{v}_{\min}) + \mathbf{v}_{\min}$. To evaluate $\Delta\text{php}(Q)$, we use $Q = \mathbf{Q}_{\max}$, because it is expected that the maximum difference occurs in $\mathrm{Q}_{\max}$ due to the HPF being directly proportional to $Q$.

We use the parameter PHH as a reference to generate the partitions on $Q$ for each hydro. N. Avanhandava, Jupiá, and Foz do Chapecó plants account for approximately 50% of the hydro installed capacity. The number of partitions chosen for these plants results in an average approximation error of less than 3%; for the other plants, we use $\mathbb{P}=\mathbb{I}$. The partitions are defined for each non-forbidden zone, and the points on $Q$ are defined based on a modification of the Douglas-Pecker algorithm [34] stated as follows:

**Algorithm 1**. (*Modified Douglas-Pecker*)

Input: number of points (NP), minimum average error (MAE), set of non-forbidden zones ( $\mathbb{I}$) and
$\mathbb{P} = [\varnothing]$;
Step 1 – For each $i \in$ $\mathbb{I}$:
Step 2 – Set $p = i$, i.e., $p = [Q_{i1}, Q_{i2}]$;

(*continued on next page*)

**Table 2**
Characteristics of hydro plants.

| Plant | PHH (%) | PHP (%) | $\Delta\text{php}(\mathrm{Q}_{\max})$ (%) | | |
|---|---|---|---|---|---|
| | | | $\gamma = 30$ | $\gamma = 60$ | $\gamma = 90$ |
| PROMISSAO | 7.37 | 3.14 | 2.57 | 2.34 | 2.28 |
| BARRA BONITA | 3.91 | 1.67 | 5.26 | 3.49 | 3.12 |
| N. AVANHANDAVA | 9.70 | 4.14 | 0 | 0 | 0 |
| JUPIÁ | 15.47 | 6.60 | 0 | 0 | 0 |
| BARIRI | 4.02 | 1.71 | 0 | 0 | 0 |
| MONJOLINHO | 2.06 | 0.88 | 3.41 | 3.38 | 3.34 |
| QUEBRA QUEIXO | 3.35 | 1.43 | 3.44 | 3.30 | 1.95 |
| SÃO JOSÉ | 1.42 | 0.61 | 0 | 0 | 0 |
| PASSO SÃO JOÃO | 2.15 | 0.92 | 4.43 | 4.37 | 4.32 |
| PASSO FUNDO | 6.31 | 2.69 | 0.10 | 0.08 | 0.66 |
| PEDRA DO CAVALO | 4.47 | 1.91 | 0.17 | 0.15 | 0.14 |
| BALBINA | 6.98 | 2.98 | 0.61 | 0.49 | 0.42 |
| GARIBALDI | 5.19 | 2.22 | 6.10 | 5.99 | 4.57 |
| FOZ DO CHAPECÓ | 23.88 | 10.19 | 2.02 | 2.01 | 2.00 |
| IBITINGA | 3.67 | 1.56 | 0 | 0 | 0 |

**Table 3**
Number of partitions for each non-forbidden zone and PWL model chosen.

| Plant | NoP for Non-Forbidden zone # | | | PWL-1 | PWL-2 |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | | |
| PROMISSAO | 1 | 1 | 1 | | X |
| BARRA BONITA | 1 | 1 | | X | |
| N. AVANHANDAVA | 1 | 1 | 1 | X | |
| JUPIÁ | 2 | | | X | |
| BARIRI | 1 | 1 | 1 | X | |
| MONJOLINHO | 1 | 1 | 1 | X | |
| QUEBRA QUEIXO | 1 | 1 | 1 | X | |
| SÃO JOSÉ | 1 | 1 | | X | |
| PASSO SÃO JOÃO | 1 | 1 | | X | |
| PASSO FUNDO | 1 | 1 | 1 | X | |
| PEDRA DO CAVALO | 1 | 1 | | X | |
| BALBINA | 1 | 1 | | X | |
| GARIBALDI | 1 | 1 | 1 | | X |
| FOZ DO CHAPECÓ | 1 | 1 | 2 | | X |
| IBITINGA | 1 | 1 | 1 | X | |

**Table 4**
Information about the optimization problems.

| Simulation Type | Ncons | Nbvar | Ncvar | RFobj | Fobj 1% | Runtime (s) | |
|---|---|---|---|---|---|---|---|
| | | | | | | Gap 1% | Gap 0.1% |
| CH | 35,936 | 2,880 | 12,529 | 454,583 | 482,986 | 6.10 | 27.79 |
| UCH | 37,376 | 3,240 | 12,529 | 454,583 | 483,160 | 8.70 | 31.34 |
| PCH | 35,096 | 3,744 | 14,257 | 454,986 | 482,253 | 103.37 | 233.88 |
| PWL | 35,864 | 3,792 | 14,593 | 455,729 | 484,516 | 33.00 | 71.15 |

(*continued*)

Step 3 – Discretize $[Q_{i1}, Q_{i2}]$ into NP = P + 1 points obtaining the set $[Q_{i1}, \ldots, Q_k, \ldots, Q_{i2}]$;
Step 4 – compute the $NA_{err}$ from equation (44);
Step 5 – If $NA_{err} \leq$ MAE: go to Step 7;
Step 6 – Else: find $k \in [Q_{i1}, \ldots, Q_k, \ldots, Q_{i2}]$ that generates the highest approximation error; insert $k$ in $p$; Go to Step 4;
Step 7 – Let the resulting set $p = [Q_1, Q_2, \ldots, Q_K]$, in which $Q_1 = Q_{i1}$, $Q_K = Q_{i2}$ and $Q_1 < Q_2 < \ldots < Q_K$. For $k = 1, \ldots, K-1$, take the interval $[Q_k, Q_{k+1}]$ and insert on

$\mathbb{P}$.
Step 8 – End For
Output: the set
$\mathbb{P}$.

Algorithm 1 can be executed *a priori* and has no impact on the computational time. To decide which PWL type must be used for each hydro, PHH and $\Delta php(\mathbf{Q}_{max})$ are analyzed. If PHH > 5% and $\Delta php(\mathbf{Q}_{max}) > 2\%$, the model chosen for the HPF is the PWL-2; otherwise, the option is PWL-1. With all these considerations, Table 3 contains the number of partitions (NoP) used in each non-forbidden zone and highlights the PWL model choose.

Finally, the system reserve is equal to 5% of the demand. The equation (60) is imposed only for the hydros with large reservoirs, and the volume target admitted is equal to 2% of variation on the initial volume, i.e.,$\mathbf{v}_h^* = 98\%\mathbf{v}_{h,0}$.

### 4.2. Results

The main results obtained from all experiments are presented in this section. Table 4 shows information about the size of optimization problems, the relaxed objective function, and runtime for different optimality gaps, where Ncons is the number of constraints of the optimization model, Ncvar is the number of continuous variables of the problem, (R)Fobj is the (relaxed) objective function value obtained with 1% gap (for the MILP). Since variations on the initial volume only impact the runtime, we present information with 60% initial volume and inflow equal to $\mathbf{Y}_1$. We observe that the PCH model is three times slower compared with the PWL model. Also, the computational effort for the UCH is relatively small if compared with the CH. On the other hand, the PWL model has significant computational effort compared with the UCH model. Moreover, the gap between the RFobj and Fobj with a 1% gap is around 6% for all models.

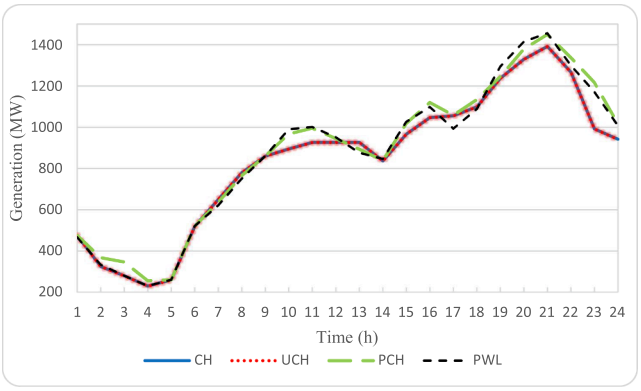
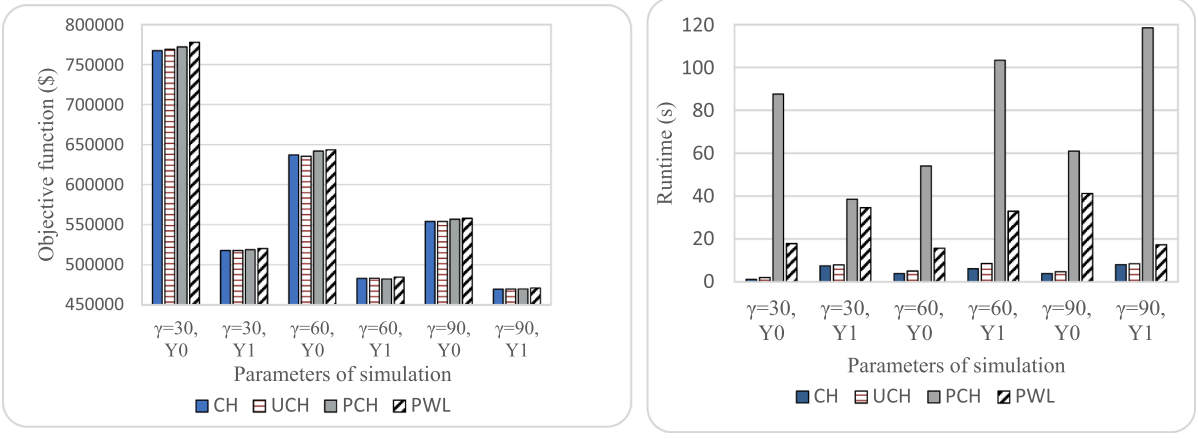



**Fig. 5.** Thermal generation in Subsystem 2 for $\gamma = 30$ and $\mathbf{Y_0}$.

For the rest of this section, all results are obtained considering an optimality gap equal to 1%. Regarding the formulation and the impact on the computational burden, the results are presented in Fig. 4. As can be seen, the CH and UCH models are solved in a few seconds. On the other hand, experiments using the PWL model consume an average of seven times more computational time compared to the CH experiment. And for the PCH model, this average increases to 23.4. The optimal cost increases with the same $\gamma$ and $\mathbf{Y}$ in the following order: CH (less expensive), UCH, PCH, and PWL. Although this does not happen in all experiments, due to the gap used, it is expected that the objective function increases as the HPF formulation gets more details.

The impact on dispatch is directly noticeable when using models with more detailed HPF. For example, in experiments with $\gamma = 30$ and




(a) – Objective function ($) (b) – Runtime (s)

**Fig. 4.** Results obtained from experiments, where (a) objective function ($), and (b) computational time (s) spent.

**Table 5**
Violations due to the operation on forbidden zones for experiments with CH model.

| Plant | $\gamma = 30$ | | $\gamma = 60$ | | $\gamma = 90$ | |
|---|---|---|---|---|---|---|
| | $Y_0$ | $Y_1$ | $Y_0$ | $Y_1$ | $Y_0$ | $Y_1$ |
| PROMISSAO | 0 | 3.29 | 0 | 0 | 6.05 | 0.09 |
| BARRA BONITA | 0 | 0 | 4.36 | 0 | 0 | 0 |
| N. AVANHANDAVA | 0 | 8.78 | 0 | 7.32 | 0 | 7.67 |
| JUPIÁ | 0 | 0 | 0 | 0 | 0 | 0 |
| BARIRI | 30.72 | 0 | 0 | 0 | 0 | 0 |
| MONJOLINHO | 0 | 0 | 0 | 0 | 0 | 0 |
| QUEBRA QUEIXO | 63.43 | 0 | 5.17 | 0 | 38.39 | 0 |
| SÃO JOSÉ | 10.72 | 0 | 0.75 | 0 | 3.50 | 0 |
| PASSO SÃO JOÃO | 92.68 | 0 | 0 | 0 | 0 | 0 |
| PASSO FUNDO | 0 | 0 | 0 | 0 | 0 | 0 |
| PEDRA DO CAVALO | 0 | 0 | 0 | 0 | 0 | 0 |
| BALBINA | 0 | 0 | 0 | 0 | 0 | 0 |
| GARIBALDI | 63.79 | 0 | 79.12 | 21.69 | 45.71 | 10.50 |
| FOZ DO CHAPECÓ | 23.64 | 0 | 7.27 | 1.72 | 0 | 1.74 |
| IBITINGA | 0 | 0 | 0 | 0 | 0 | 0 |

**Table 6**
Violations due to the operation on forbidden zones for experiments with UCH model.

| Plant | $\gamma = 30$ | | $\gamma = 60$ | $\gamma = 90$ | | |
|---|---|---|---|---|---|---|
| | $Y_0$ | $Y_1$ | $Y_0$ | $Y_1$ | $Y_0$ | $Y_1$ |
| PROMISSAO | 1.96 | 0 | 0 | 3.20 | 3.68 | 5.48 |
| BARRA BONITA | 0 | 0 | 1.67 | 0 | 0 | 0 |
| N. AVANHANDAVA | 0 | 4.26 | 0 | 2.41 | 0 | 2.44 |
| JUPIÁ | 0 | 0 | 0 | 0 | 0 | 0 |
| BARIRI | 34.80 | 0 | 0 | 0 | 0 | 0 |
| MONJOLINHO | 0 | 0 | 0 | 0 | 0 | 0 |
| QUEBRA QUEIXO | 18.09 | 0 | 13.70 | 0 | 12.13 | 0 |
| SÃO JOSÉ | 0 | 0 | 4.23 | 0 | 3.34 | 0 |
| PASSO SÃO JOÃO | 41.05 | 0 | 0 | 0 | 0 | 0 |
| PASSO FUNDO | 0 | 0 | 0 | 0 | 0 | 0 |
| PEDRA DO CAVALO | 0 | 0 | 0 | 0 | 0 | 0 |
| BALBINA | 0 | 0 | 0 | 0 | 0 | 0 |
| GARIBALDI | 14.79 | 0 | 51.35 | 10.46 | 29.17 | 3.54 |
| FOZ DO CHAPECÓ | 25.74 | 0 | 10.59 | 1.53 | 0 | 0 |
| IBITINGA | 0 | 0 | 0 | 0 | 0 | 0 |

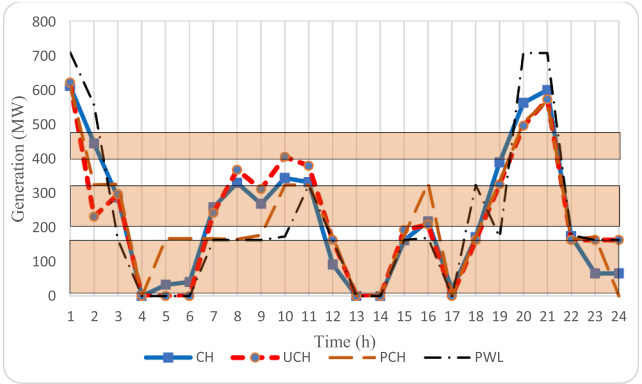



**Fig. 6.** Scheduled generation of Foz do Chapecó plant.

$\mathbf{Y}_0$, when analyzing the total of thermal generation of subsystem 2 [32], presented in Fig. 5, we notice an increase in the generation for the PCH and PWL models of, respectively, 4.32 and 3.01%, when compared with the dispatch obtained from CH model.

The results in Tables 5 and 6 present the violations on forbidden zones. Note that this only occurs on the CH and UCH models. The results can be interpreted in the following way. First, consider $\gamma = 30$ and $\mathbf{Y}_0$. In Table 5, for the Garibaldi, considering the power scheduled for each hour of the planning horizon obtained through the optimization problem, 63.79% occur in forbidden zones. For the UCH model, on the other hand, the generation in forbidden zones is 14.79%. This number, combined with the PHH and the total scheduled for this plant, can help the ISO to assess the feasibility of the generation scheduled.

**Table 7**
Results for different parameters of PWL formulation.

| Simulation # | PWL-# | | NoP for each non-forbidden zone | Fobj ($) | Runtime (s) | OHPFe (%) |
|---|---|---|---|---|---|---|
| | 1 | 2 | | | | |
| 1 | X | | 3 | 483,138 | 47.70 | 5.41 |
| 2 | | X | 3 | 484,024 | 126.92 | 5.19 |
| 3 | X | | 5 | 484,515 | 43.28 | 5.24 |
| 4 | | X | 5 | 482,368 | 325.70 | 4.79 |
| 5 | X | | 8 | 483,122 | 75.31 | 4.74 |
| 6 | | X | 8 | 483,744 | 600 | 4.62 |
| Ref | – | – | – | 484,516 | 33.00 | 5.64 |

For illustrative purposes, consider Fig. 6 that presents the Foz do Chapecó plant generating scheduling over the planning horizon. The shadow rectangles are highlighting the forbidden zones. This plant possesses PHH = 23.88%. Note that almost 25% of the generation for the CH and UCH operate in the forbidden zones. Fig. 6 shows that the CH model operation occurs in forbidden zones within several hours of the planning horizon. On the other hand, in the UCH model, this aspect occurs less frequently.

We use Table 7 to show the impact of the volume effects in the hydro plants. Also, we have increased the number of partitions for each non-forbidden zone (equally spaced) to have an HPF model closer to the nonlinear function. For comparison, we include the measure of overall HPF error (OHPFe), which is the relative error over the planning horizon for all hydro plants related with the total hydro generation, and the results obtained for the reference (Ref) simulation, where for the PWL model we used the mixed strategy presented on Table 3. All simulations are performed considering 1% of the gap, $\gamma = 60$, $\mathbf{Y}_1$, and time-limit of 600 s.

Notice that the PWL formulation can achieve reasonable OHPFe values, although the increase of NoP is the price to pay for obtaining better precision. Also, the inclusion of the volume effects on HPF has a significant impact on the computational effort. As can be seen, the different formulations on the HPF can directly impact the NCHTUC problem. For these experiments, a relative number of hydro plants showed violations on the scheduling for the CH and UCH model. Of course, if the total amount of violation is relatively high, the PCH and PWL models can fix this problem.

On the other hand, the UCH formulation can be interesting in cases where the violations are minor. The PWL formulation is attractive if a high level of accuracy is required for the HPF, but many binary variables are required depending on the precision, and therefore, a higher computational effort is required. Also, it is possible to impose a particular formulation for each hydro plant of the system, but this is out of this work scope.

## 5. Conclusions and future work

Due to the nonlinear characteristics of the HPF, the NCHTUC problem is usually approached by an MINLP, and, for large-scale systems, it is almost impossible to solve this problem in a reasonable time. Therefore, approximations are imposed on the HPF, where the plant-based formulation is the approach used in this work. If the system is hydro predominant, models with great details for the HPF are attractive. In this sense, this work proposes different formulations for the HPF, aiming to insert the nonlinear characteristics and the forbidden zones of the GUs, where the PWL formulation makes it possible to control the approximation error. The experiments are performed on a modified version of the IEEE-118 system, where hydros were included. From the results, we

observe that the PCH and PWL models are an attractive alternative to represent the HPF since both models can handle precisely the operation in forbidden zones. The increase in computational time is an issue that needs more investigation in large-size systems, although decomposition techniques and parallel processing are natural options in this context. Also, depending on the accuracy required, a particular HPF model can be chosen, enabling an NCHTUC where each hydro has a specific type of formulation presented in this work. Future works might include the effects of spillage, assess the impact of different representations of PWL models, i.e., the logarithmic and SOS-2 representation, and apply the different HPF models presented on a large-scale system.

## CRediT authorship contribution statement

**K.V. Santos:** Conceptualization, Methodology, Software, Writing – original draft. **E.C. Finardi:** Conceptualization, Visualization, Investigation, Supervision, Writing – review & editing, Project administration, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

The authors thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and Norte Energia, via R&D Project registered with the PD-07427-0318/2018 code in ANEEL (Agência Nacional de Energia Elétrica), for their financial support.

## References

[1] S. Stoft, Power system economics: designing markets for electricity. Piscataway, NJ: New York: IEEE Press, Wiley-Interscience, 2002.
[2] P. Cramton, Electricity market design, Oxford Review of Economic Policy, vol. 33, no 4, p. 589–612, nov. 2017, doi: 10.1093/oxrep/grx041.
[3] W. Van Ackooij, I. Danti Lopez, A. Frangioni, F. Lacalandra, E.M. Tahanan, Large-scale unit commitment under uncertainty: an updated literature survey, Ann Oper Res, 271 (1), 2018, p. 11–85, doi: 10.1007/s10479-018-3003-z.
[4] E. C. Finardi, R. D. Lobato, V. L. de Matos, C. Sagastizábal, e A. Tomasgard, Stochastic hydrothermal unit commitment via multi-level scenario trees and bundle regularization, Optim Eng, vol. 21, no 2, p. 393–426, jun. 2020, doi: 10.1007/s11081-019-09448-z.
[5] Colonetti B, Finardi EEC. Stochastic hydrothermal unit commitment models via stabilized benders decomposition. Electr. Eng. 2021. https://doi.org/10.1007/s00202-020-01206-0.
[6] S.-E. Razavi, A. Esmaeel Nezhad, H. Mavalizadeh, F. Raeisi, E.A. Ahmadi, Robust hydrothermal unit commitment: A mixed-integer linear framework, Energy, vol. 165, p. 593–602, dec. 2018, doi: 10.1016/j.energy.2018.09.199.
[7] J. Kong, H. I. Skjelbred, e O. B. Fosso, An overview on formulations and optimization methods for the unit-based short-term hydro scheduling problem, Electric Power Systems Research, vol. 178, p. 106027, jan. 2020, doi: 10.1016/j.epsr.2019.106027.
[8] A. L. Diniz e M. E. P. Maceira, A Four-Dimensional Model of Hydro Generation for the Short-Term Hydrothermal Dispatch Problem Considering Head and Spillage Effects, IEEE Trans. Power Syst., vol. 23, no 3, p. 1298–1308, aug. 2008, doi: 10.1109/TPWRS.2008.922253.
[9] J. P. S. Catalao, S. J. P. Mariano, V. M. F. Mendes, e L. A. F. Ferreira, Scheduling of Head-Sensitive Cascaded Hydro Systems: A Nonlinear Approach, IEEE Trans. Power Syst., vol. 24, no 1, p. 337–346, feb. 2009, doi: 10.1109/TPWRS.2008.2005708.
[10] J. P. S. Catalão, H. M. I. Pousinho, e V. M. F. Mendes, Mixed-integer nonlinear approach for the optimal scheduling of a head-dependent hydro chain, Electric Power Systems Research, vol. 80, no 8, p. 935–942, aug. 2010, doi: 10.1016/j.epsr.2009.12.015.
[11] J. P. S. Catalão, H. M. I. Pousinho, e V. M. F. Mendes, Hydro energy systems management in Portugal: Profit-based evaluation of a mixed-integer nonlinear approach, Energy, vol. 36, no 1, p. 500–507, jan. 2011, doi: 10.1016/j.energy.2010.10.014.
[12] A. Hamann, G. Hug, e S. Rosinski, Real-Time Optimization of the Mid-Columbia Hydropower System, IEEE Trans. Power Syst., vol. 32, no 1, p. 157–165, jan. 2017, doi: 10.1109/TPWRS.2016.2550490.
[13] A. L. Diniz e T. M. Souza, Short-Term Hydrothermal Dispatch With River-Level and Routing Constraints, IEEE Trans. Power Syst., vol. 29, no 5, p. 2427–2435, sep. 2014, doi: 10.1109/TPWRS.2014.2300755.
[14] L. S. M. Guedes, P. de Mendonca Maia, A. C. Lisboa, D. A. G. Vieira, e R. R. Saldanha, A Unit Commitment Algorithm and a Compact MILP Model for Short-Term Hydro-Power Generation Scheduling, IEEE Trans. Power Syst., vol. 32, no 5, p. 3381–3390, sep. 2017, doi: 10.1109/TPWRS.2016.2641390.
[15] S. Seguin, P. Cote, e C. Audet, Self-Scheduling Short-Term Unit Commitment and Loading Problem, IEEE Trans. Power Syst., vol. 31, no 1, p. 133–142, jan. 2016, doi: 10.1109/TPWRS.2014.2383911.
[16] A. Marchand, M. Gendreau, M. Blais, e G. Emiel, Fast Near-Optimal Heuristic for the Short-Term Hydro-Generation Planning Problem, IEEE Trans. Power Syst., vol. 33, no 1, p. 227–235, jan. 2018, doi: 10.1109/TPWRS.2017.2696438.
[17] K. Koch, B. Hiller, M. E. Pfetsch, e L. Schewe, Evaluating gas network capacities, SIAM, 2015.
[18] Balakrishnan A, Graves ESC. A composite algorithm for a concave-cost network flow problem. Networks 1989:175–202.
[19] B. H. Brito, E. C. Finardi, e F. Y. K. Takigawa, Mixed-integer nonseparable piecewise linear models for the hydropower production function in the Unit Commitment problem, Electric Power Systems Research, vol. 182, p. 106234, may 2020, doi: 10.1016/j.epsr.2020.106234.
[20] A. Borghetti, C. D'Ambrosio, A. Lodi, e S. Martello, An MILP Approach for Short-Term Hydro Scheduling and Unit Commitment With Head-Dependent Reservoir, IEEE Trans. Power Syst., vol. 23, no 3, p. 1115–1124, aug. 2008, doi: 10.1109/TPWRS.2008.926704.
[21] B. Tong, Q. Zhai, e X. Guan, An MILP Based Formulation for Short-Term Hydro Generation Scheduling With Analysis of the Linearization Effects on Solution Feasibility, IEEE Trans. Power Syst., vol. 28, no 4, p. 3588–3599, nov. 2013, doi: 10.1109/TPWRS.2013.2274286.
[22] X. Li, T. Li, J. Wei, G. Wang, e W. W.-G. Yeh, Hydro Unit Commitment via Mixed Integer Linear Programming: A Case Study of the Three Gorges Project, China, IEEE Trans. Power Syst., vol. 29, no 3, p. 1232–1241, may 2014, doi: 10.1109/TPWRS.2013.2288933.
[23] Jiangtao Jia e Xiaohong Guan, MILP formulation for short-term scheduling of cascaded reservoirs with head effects, in 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), Deng Feng, China, aug. 2011, p. 4061–4064, doi: 10.1109/AIMSEC.2011.6009930.
[24] Gurobi Optimizer Reference Manual. 2020, [Online]. Available in: http://www.gurobi.com.
[25] H. I. Skjelbred, J. Kong, e O. B. Fosso, Dynamic incorporation of nonlinearity into MILP formulation for short-term hydro scheduling, International Journal of Electrical Power & Energy Systems, vol. 116, p. 105530, mar. 2020, doi: 10.1016/j.ijepes.2019.105530.
[26] I. Guisández e J. I. Pérez-Díaz, Mixed integer linear programming formulations for the hydro production function in a unit-based short-term scheduling problem, International Journal of Electrical Power & Energy Systems, vol. 128, p. 106747, jun. 2021, doi: 10.1016/j.ijepes.2020.106747.
[27] E. C. Finardi e E. L. da Silva, Solving the Hydro Unit Commitment Problem via Dual Decomposition and Sequential Quadratic Programming, IEEE Transactions on Power Systems, vol. 21, no 2, p. 835–844, may 2006, doi: 10.1109/TPWRS.2006.873121.
[28] J. P. Vielma, S. Ahmed, e G. Nemhauser, Mixed-Integer Models for Nonseparable Piecewise-Linear Optimization: Unifying Framework and Extensions, Operations Research, vol. 58, no 2, p. 303–315, apr. 2010, doi: 10.1287/opre.1090.0721.
[30] Wallace SW, Fleten ES-E. Stochastic Programming Models in Energy, in Handbooks in Operations Research and Management Science. Elsevier 2003;10:637–77.
[31] K. W. Hedman, M. C. Ferris, R. P. O'Neill, E. B. Fisher, e S. S. Oren, Co-Optimization of Generation Unit Commitment and Transmission Switching With N-1 Reliability, IEEE Transactions on Power Systems, vol. 25, no 2, p. 1052–1063, may 2010, doi: 10.1109/TPWRS.2009.2037232.
[32] IEEE 118-bus data. 2020, [Online]. Available in: http://motor.ece.iit.edu/data/.
[33] Douglas DH, Peucker ETK. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Can. Cartogr. 1973:112–22.
[34] Jilin C, Min Z, Zhonghua G, Weijiang Q, Yong C, Weixi W. The Application of Douglas-Peucker Algorithm in Collaborative System for Power Grid Operation Mode Calculation. MATEC Web Conf. 2018;175:03041. https://doi.org/10.1051/matecconf/201817503041.