

## Article

# Continuous Piecewise Linear Approximation of Plant-Based Hydro Production Function for Generation Scheduling Problems

David Lucas dos Santos Abreu <sup>1</sup>  and Erlon Cristian Finardi <sup>1,2,\*</sup> 

<sup>1</sup> Electric Energy Systems Planning Laboratory (LabPlan), Federal University of Santa Catarina, Florianópolis 88040-900, Brazil; dlbr@live.com

<sup>2</sup> Institute of Systems and Computer Engineering, Research and Development of Brazil (INESC P&D Brasil), Santos 11055-300, Brazil

\* Correspondence: erlon.finardi@ufsc.br

**Abstract:** An essential challenge in generation scheduling (GS) problems of hydrothermal power systems is the inclusion of adequate modeling of the hydroelectric production function (HPF). The HPF is a nonlinear and nonconvex function that depends on the head and turbined outflow. Although the hydropower plants have multiple generating units (GUs), due to a series of complexities, the most attractive modeling practice is to represent one HPF per plant, i.e., a single function is built for representing the plant generation instead of the generation of each GU. Furthermore, due to the computation time constraints and representation of nonlinearities, the HPF must be given by a piecewise linear (PWL) model. This paper presented some continuous PWL models to include the HPF per plant in GS problems of hydrothermal systems. Depending on the type of application, the framework allows a choice between the concave PWL for HPF modeled with one or two variables and the nonconvex (more accurate) PWL for HPF dependent only on the turbined outflow. Basically, in both PWL models, offline, mixed-integer linear (or quadratic) programming techniques are used with an optimized pre-selection of the original HPF dataset obtained through the Ramer-Douglas-Peucker algorithm. As a highlight, the framework allows the control of the number of hyperplanes and, consequently, the number of variables and constraints of the PWL model. To this end, we offer two possibilities: (i) minimizing the error for a fixed number of hyperplanes, or (ii) minimizing the number of hyperplanes for a given error. We assessed the performance of the proposed framework using data from two large hydropower plants of the Brazilian system. The first has 3568 MW distributed in 50 Bulb-type GUs and operates as a run-of-river hydro plant. In turn, the second, which can vary the reservoir volume by up to 1000 hm<sup>3</sup>, possesses 1140 MW distributed in three Francis-type units. The results showed a variation from 0.040% to 1.583% in terms of mean absolute error and 0.306% to 6.356% regarding the maximum absolute error even with few approximations.

**Keywords:** hydro production function; piecewise linear model; mixed-integer linear programming



**Citation:** dos Santos Abreu, D.L.; Finardi, E.C. Continuous Piecewise Linear Approximation of Plant-Based Hydro Production Function for Generation Scheduling Problems. *Energies* **2022**, *15*, 1699. <https://doi.org/10.3390/en15051699>

Academic Editor: Helena M. Ramos

Received: 2 February 2022

Accepted: 22 February 2022

Published: 24 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The electricity industry in the world has gone through significant transitions regarding the digitalization of electric networks, management and automation of industries, expansion in the use of renewable sources, and improvement of decentralized energy generation technologies. This process is synthesized in the so-called three Ds: decarbonization, digitization, and decentralization. In this context, Brazil, for example, has a privileged position, having, in 2020, about 83% of the electrical power system made up of renewable energy resources. However, it is worth noting that a considerable amount of these resources, about 65.2%, comes from hydroelectricity, whose expansion is limited due to solid environmental constraints. Thus, it is essential to optimize the operation of this type of resource, which is performed by operators and generation companies worldwide, mainly with the support of long, medium, and short-term generation scheduling (GS) models [1]. The mathematical

characteristics of these models are intrinsically linked to the market design employed in a given electrical system. In systems whose design adopts a loose-pool dispatch, the GS models are tools primarily used to maximize the expected revenue associated with energy trading in spot and future markets [2,3]. On the other hand, in the case of markets with centralized dispatch, i.e., the tight-pool, an independent system operator employs a chain of GS models to minimize the expected operating cost, considering some risk-measure [4–7].

Regardless of the market design, the representation of the hydroelectric production function (HPF) is a vital modeling aspect in each GS model. Briefly, the HPF relates the power output with the turbined outflow, net head, and efficiency. In turn, the generating unit (GU) efficiency is a function of the turbined outflow and net head. As a result, the HPF is usually represented by a two-variable, nonlinear, and nonconvex function. In addition to the representation by GU, the HPF can represent the plant generation by aggregating the GUs into an equivalent GU [6]. In long and medium-term GS models, the inclusion of uncertainties for risk management via, for example, the stochastic programming model is a priority [4]. In this case, the dimensionality and mathematical characteristics of the HPF must be controlled so that the models can efficiently include the uncertainties.

Furthermore, with the increasing insertion of wind generation and run-of-river hydroelectric plants, short-term GC models also include uncertainties in their modeling, demanding the same type of requirement as the more extended horizon models. In this context, the plant-based HPF is advantageous because the high dimensionality introduced by the individualized representation considerably increases the GS models' computational burden. Although the plant-based representation is the most feasible option, other aspects that burden the solving strategies of GS models are the nonlinearities and nonconvexities inherent to the HPF. In this sense, an alternative that has proven to be very efficient is the HPF representation by a piecewise linear function (PWL), which allows GS problems to be effectively solved via state-of-the-art linear programming (LP) and mixed-integer linear programming (MILP) software.

The literature is very vast in the representation of HPF in GS problems. Thus, to support the developments proposed in this paper, in the sequence, we focused on more recent works that make use of the HPF aggregated representation, commonly called plant-based models. Thus, an efficient numerical optimization method to obtain an HPF is proposed for cascading reservoirs in long-term GS problems [5]. The HPF depends on the average stored volume and the turbined outflow in this work. Despite the good accuracy of this method, the nonlinear representation of the hydraulic performance is not considered as in [6]. Still, in [6], the HPF was represented by a PWL model, which depends on the volume and the turbined outflow obtained by an algorithm incorporating convex hull (CH) techniques. This model is distinguished by using the nonlinear HPF instead of constant productivity models [8,9], whose HPF depends solely on the turbined outflow (therefore, it does not include the effects of the net head variations). Another common approach is the use of univariate PWL. In such a case, the gross head is fixed, and the linearization is accomplished for an HPF that depends on the turbined outflow [7]. The main advantage of this method is the implementation's simplicity and suitable computational performance. However, the non-consideration, simultaneously, of the net head and turbined outflow effects results in a flawed correspondence between the PWL and original HPF. Another alternative that can be used is the direct representation of the HPF by a set of concave functions dependent on the turbined outflow, according to [9,10]. However, this type of representation notably increases the number of constraints and variables in the problem.

The plant-based HPF is also widely used in the context of short-term GS models, whether in purely hydro systems [11–13], hydrothermal systems [14–18], or hydrothermal systems interconnected with renewable sources such as solar and wind [14–17]. In [15], it PWL is obtained as a function of the turbined outflow and gross head through CH techniques. It is central to highlight that the resulting model obtained by the CH is concave. Therefore, the PLW can be included in the GS problems without adding binary variables, significantly favoring the computational performance. However, these approximations

often overestimate the function, mainly in the nonconcave regions of the HPF, usually requiring coefficients adjustment or hybrid linearization strategies [19]. The models proposed in [14–17] represent the HPF as a quadratic function dependent on the volume and the turbined outflow. Although the product between outflow and the net head has been included in this case, the use of quadratic functions inevitably leads to mixed-integer quadratic programming models (MIQP), which computationally burdens the resolution of the problem. The model proposed in [16] used a nonlinear HPF depending on gross head and turbined flow. This approach requires the use of nonlinear programming (NLP) techniques that, although resulting in accurate solutions, are impractical for use in large-scale systems. Additionally, in some works, as in [18,20], the linearization is based on SOS2 sets. The idea is that the power generation, gross head, and turbined outflow can be expressed as a convex combination of SOS2 sets.

Some studies have relied on linearization based on the logarithmic aggregation convex combination (LACC) model [19,21]. According to the current literature, the LACC model represents one of the best options for representing high-precision PWL functions. However, the number of equations associated with LACC is intrinsically linked to the number of data points employed in the PWL approximation. This issue leads to incorporating more binary variables in CG models, further increasing the computational effort. In [21], for handling this negative issue, the LACC is employed only in the nonconvex domain of the HPF, and the concave part is linearized by using CH techniques. Nevertheless, the overall PWL approach can often supply unnecessary equations for a given precision due to the lack of “controllability,” which is a key point we offer in this paper.

The primary purpose of this paper was to present a computational framework based on the most recent contributions developed for PWL fitting. A relevant particularity is that an analytical model is not available for the plant-based HPF, unlike the individualized case, unless some simplifications are adopted [6,22]. Thus, for the construction of the plant-based PWL, it must be considered that only a discrete set of points are provided, given by the triples (turbined outflow, gross head, and power generation) for the bivariate case where outflow and gross head are input parameters, or (outflow and power generation) for the univariate case where the gross head is fixed. In particular, the power generation values can only be obtained by solving a mixed-integer nonlinear optimization problem (MINLP) since the maximum power generation of the plant depends on which combination of GUs is used in the optimal dispatch [23].

Recently, two studies presented significant contributions in PWL fitting [24,25], which avoids the use of nonlinear constraints necessary to preserve the continuity of the fitted model. Precisely, those works use big-M-type constructions to enforce continuity through a set of linear constraints. This paper aimed to use these formulations with modifications and contributions in the plant-based HPF.

Few papers in the literature propose a representation of the plant-based HPF with a PWL function obtained through a refined control of the number of hyperplanes or a pre-established linearization error. Therefore, the major contribution of this work is the “controllability” aspect. Our main objective was to present a computational framework based on the most recent contributions developed in PWL and apply it to the plant-based HPF. We also propose using the Ramer-Douglas-Peucker algorithm (RDP) [26,27] for the optimal selection of points to be used in the linearization process. Moreover, this approach also selects the nonlinear region of the HPF, which was our region of interest.

Firstly, this work showed the nonlinear HPF for the individualized case (i.e., GU), and then it built the optimization model based on MINLP to extract the plant-based HPF. The models then linearized the nonlinear HPF. In the proposed framework, the control is achieved in two ways: (i) minimizing the number of hyperplanes for a given error and (ii) minimizing the error for a fixed number of hyperplanes. By stipulating a given acceptable error in the linearization, it is possible to obtain the minimum amount of associated hyperplanes, thus limiting the number of variables in the problem and improving the computational performance. As a result of this, we can establish, for example,

acceptable errors for different hydropower plants according to their importance in the system (e.g., power or storage capacity). The predictability of computational performance is improved if we choose to fix the number of hyperplanes, thus knowing the number of variables and constraints of the model in advance. The PWL models can still be restricted to produce nonconvex and concave PWL functions. Although more accurate, the nonconvex model requests binary variables, hindering use in large-scale problems. In this work, the nonconvex PWL was constrained to the univariate case, thus being more suitable for plants with lower capacity or simplified representation in the final stages of the planning horizon of a given GS model.

On the other hand, the concave PWL eliminates binary variables to represent the plant-based HPF. This paper used the concave approach in both univariate and bivariate cases. Subsequently, we presented two different methods of partitioning the HPF domain: (i) the most used in the literature which is based on uniform discretization, i.e., the domain is subdivided into a grid of equidistantly spaced points, and (ii) in which an optimal partition of the domain is performed using the RDP algorithm which, in general terms, searches for a curve with a higher concentration of points in regions of pronounced nonlinearity of the HPF. With the appropriate treatment of the data, the models proposed in this work then control the accuracy and the resulting computational effort in the GS models.

The remainder of this paper is organized as follows: Section 2 provides details on obtaining the plant-based HPF. Section 3 presents the proposed framework for the linearization of the HPF. Section 4 presents the use of the RDP algorithm for the optimized pre-selection of data. Numerical results and discussions are provided in Section 5. Finally, conclusions are discussed in Section 6.

## 2. The Plant-Based Hydro Production Function

To present the plant-based HPF formulation, it is important first to detail the individual model, which is given by:

$$g_j = 0.00981 \cdot ge_j \cdot w_j \cdot h_j \quad (1)$$

In (1),  $g_j$  is the power generation of the GU  $j$  (MW),  $ge_j$  is the global efficiency of the GU  $j$ ,  $w_j$  is the turbined outflow of the GU  $j$  ( $\text{m}^3/\text{s}$ ), and  $h_j$  is the net head of the GU  $j$  (m). The net head is the difference between the gross head,  $gh$ , and the hydraulic losses associated with the turbined outflow. By definition,  $gh$  is given by the difference between the forebay and the tailrace levels, which depend in most cases on the stored volume ( $v$ ) and the total outflow ( $d$ ) of the reservoir, respectively, as follows:

$$gh = fbl(v) - trl(d) \quad (2)$$

In (2),  $fbl(v)$  is the forebay level (m) and  $trl(d)$  is the tailrace level (m). The total outflow  $d$  is given by  $q + s$ , where  $q$  is the plant turbined outflow. Mathematically, the equation  $q = \sum_{j=1}^N w_j$  represents the turbined outflow in the plant ( $\text{m}^3/\text{s}$ ), and  $s$  is the spillage ( $\text{m}^3/\text{s}$ ).

We momentarily put aside  $fbl(v)$  and  $trl(d)$  to write the HPF model with only two variables. The dependence of  $gh$  with the forebay and tailrace levels is directly expressed in GS models as equality constraints or even as PWL. In that way, the net head is expressed as:

$$h_j = gh - \mathbf{K}_j w_j^2 \quad (3)$$

In (3),  $\mathbf{K}_j$  is the constant used to represent the hydraulic loss function in the GU  $j$  ( $\text{s}^2/\text{m}^5$ ), which considers that the GU has an individual penstock. When this is not the case, (3) also depends on the turbined outflow of other GUs of the same plant. Although modeling (3) is the most common, the framework presented in this paper does not require this restriction. On the other hand,  $ge_j$  in (1) includes several kinds of turbine and generator efficiencies (hydraulic, mechanical, and electrical). Based on field tests and hill-curve data from GU, a nonlinear function that depends mostly on  $w_j$  and  $h_j$  can express the global

efficiency. We assume that the following polynomial function is available (although a different function can be used):

$$ge_j = \mathbf{E}_{0j} + \mathbf{E}_{1j}w_j + \mathbf{E}_{2j}h_j + \mathbf{E}_{3j}w_j^2 + \mathbf{E}_{4j}h_j^2 \tag{4}$$

Above in (4),  $\mathbf{E}_{kj}$  is the  $k$ th constant in the global efficiency function of the GU  $j$ . By including (2) and (3) in (1), the unit-based HPF final expression is presented below.

$$g_j = 0.00981[\mathbf{E}_{0j} + \mathbf{E}_{1j}w_j + \mathbf{E}_{2j}(gh - \mathbf{K}_jw_j^2) + \mathbf{E}_{3j}w_j^2 + \mathbf{E}_{4j}(gh - \mathbf{K}_jw_j^2)^2]w_j(gh - \mathbf{K}_jw_j^2) \tag{5}$$

As it can be seen, the expression (5) is written as a function of  $gh$  and  $w_j$  because (3) only depends on  $w_j$ . If this is not the case, then (5) must be written as a function of  $h_j$  and  $w_j$ .

Once the individual model is presented, the next step is to detail the plant-based model, the focus of this work. To do so, initially, consider that a given plant has  $\mathbf{N}$  GUs. To facilitate the presentation, also consider that each GU has only one operating zone, given by the range of the turbined outflow ( $W_j^{\min}$ – $W_j^{\max}$ ). Thus, to obtain the dimensional reduction aimed by the plant-based model, the  $\mathbf{N}$  individual HPF (5) must be replaced by a function that depends on  $gh$  and the plant turbined outflow  $q$  instead of  $w_j$ . As previously mentioned, an analytical and precise plant-based HPF is not available since the hydraulic loss and the global efficiency of the plant depend on the number of dispatched GUs for a given operating state (i.e.,  $gh$  and  $q$ ). The alternative is to obtain a dataset of triples ( $\mathbf{PG}_{kl}$ ,  $\mathbf{Q}_k$ ,  $\mathbf{GH}_l$ ) based on the solution of the MINLP presented below. The MINLP considers as input the parameters the value of  $\mathbf{Q}_k$ , for  $k = 1, \dots, \mathbf{K}$  defined in ( $\mathbf{Q}_j^{\min}$ – $\mathbf{Q}_j^{\max}$ ) and  $l = 1, \dots, \mathbf{L}$  values of  $\mathbf{GH}_l$  in the range ( $\mathbf{HB}^{\min}$ – $\mathbf{HB}^{\max}$ ).

$$\mathbf{PG}_{kl} = \text{maximize } \sum_{j=1}^{\mathbf{N}} g_j \tag{6}$$

$$\text{s.t. : } \sum_{j=1}^{\mathbf{N}} w_j = \mathbf{Q}_k, \forall j = 1, \dots, \mathbf{N} \tag{7}$$

$$g_j - 0.00981 \begin{bmatrix} \mathbf{E}_{0j} + \mathbf{E}_{1j}w_j + \mathbf{E}_{2j}(\mathbf{GH}_l - \mathbf{K}_jw_j^2) \\ + \mathbf{E}_{3j}w_j(\mathbf{GH}_l - \mathbf{K}_jw_j^2) + \mathbf{E}_{4j}w_j^2 \\ + \mathbf{E}_{5j}(\mathbf{GH}_l - \mathbf{K}_jw_j^2)^2 \end{bmatrix} (\mathbf{GH}_l - \mathbf{K}_jw_j^2)w_j = 0, \forall j = 1, \dots, \mathbf{N} \tag{8}$$

$$W_j^{\min}u_j \leq w_j \leq W_j^{\max}u_j, \forall j = 1, \dots, \mathbf{N} \tag{9}$$

$$u_j \in \{0, 1\}. \tag{10}$$

In this optimization problem,  $k$  is the index associated with the plant turbined outflow  $\mathbf{Q}_k$ ( $\text{m}^3/\text{s}$ ),  $l$  is the index associated with the gross head  $\mathbf{GH}_l$  (m),  $\mathbf{PG}_{kl}$  is the power generation of the plant in ( $\mathbf{Q}_k$ ,  $\mathbf{GH}_l$ ), and  $u_j$  is the binary indicating if the GU  $j$  is online ( $u_j = 1$ ) or offline ( $u_j = 0$ ). Notice that we are using uppercase bold notation for fixed values. Thus,  $\mathbf{K} \times \mathbf{L}$  optimization problems are solved, and, therefore, we obtain  $\mathbf{K} \times \mathbf{L}$  values of  $\mathbf{PG}_{kl}$ .

The objective function (6) maximizes the power generation by the plant. The constraint (7) delimits that the sum of the turbined outflow of the GUs is equal to the turbined outflow of the plant. The constraint (8) is the nonlinear nonconvex HPF related to the UG  $j$ . Finally, (9) defines the limits of the turbined outflow of each GU, and (10) represents the integrality constraints. The (6)–(10) formulation describes the bivariate case since the HPF depends on  $\mathbf{Q}_k$  and  $\mathbf{GH}_l$ . Now consider the univariate case, where the gross head is fixed in  $\mathbf{HB}_0$ , so we have a form dataset ( $\mathbf{PG}_k$ ,  $\mathbf{Q}_k$ ). The MINLP is now rewritten as:

$$\mathbf{PG}_k = \text{maximize } \sum_{j=1}^{\mathbf{N}} g_j \tag{11}$$

$$\text{s.t. : } \sum_{j=1}^N w_j = \mathbf{Q}_k, \forall j = 1, \dots, N \quad (12)$$

$$g_j - 0.00981 \begin{bmatrix} \mathbf{E}_{0j} + \mathbf{E}_{1j}w_j + \mathbf{E}_{2j}(\mathbf{GH}_0 - \mathbf{K}_jw_j^2) \\ + \mathbf{E}_{3j}w_j(\mathbf{GH}_0 - \mathbf{K}_jw_j^2) + \mathbf{E}_{4j}w_j^2 \\ + \mathbf{E}_{5j}(\mathbf{GH}_0 - \mathbf{K}_jw_j^2)^2 \end{bmatrix} (\mathbf{GH}_0 - \mathbf{K}_jw_j^2)w_j = 0, \forall j = 1, \dots, N \quad (13)$$

$$\mathbf{W}_j^{\min} u_j \leq w_j \leq \mathbf{W}_j^{\max} u_j, \forall j = 1, \dots, N \quad (14)$$

$$u_j \in \{0, 1\} \quad (15)$$

The accuracy of the PWL depends on the number of points  $\mathbf{K} \times \mathbf{L}$ , such as  $k = 1, \dots, \mathbf{K}$  and  $l = 1, \dots, \mathbf{L}$  ( $\mathbf{L} = 1$  in the univariate case) points used in the proposed framework. In general, the quality of a PWL fitting increases with the increase in  $\mathbf{K}$  and  $\mathbf{L}$ . However, a fair value of the total points is not simple to determine, in particular, due to the nonconvexity of the HPF. Although the dataset obtained in the Formulations (6)–(10) and (11)–(15) can be used in the algorithms presented in Section 3, it is possible to perform an optimized selection of these data using the Ramer-Douglas-Peucker (RDP) algorithm, as is shown in Section 4. The RDP increases the computational performance, as it is possible to find a PWL model with equivalent precision using fewer points.

### 3. Piecewise Linear Models

The construction of the PWL models proposed in this paper was carried out using the methodologies [19,20] based on solving MILP or MIQP optimization problems. Thus, according to Section 2, consider initially that  $(\mathbf{PG}_{kl}, \mathbf{Q}_k, \mathbf{GH}_l)$  or  $(\mathbf{PG}_k, \mathbf{Q}_k)$  is available such that  $k = 1, \dots, \mathbf{K}$  e  $l = 1, \dots, \mathbf{L}$ . Thus, we sought to determine a continuous PWL that best approximates this set of points, minimizing an error established by an  $o$ -type norm. Considering formulations as parametric models, that is, the objective is to obtain coefficients that result in a set of multiple-choice hyperplanes [28], the objective function of the PWL model optimization problem can be described for the univariate HPF, as follows:

$$\min \Theta = \sum_{k=1}^{\mathbf{K}} \left| c^1_H \mathbf{Q}_k + d_H - \mathbf{PG}_k \right|^o \quad (16)$$

In (16),  $c^1_H$  is the slope or gradient 1 associated with  $\mathbf{Q}_k$  in the hyperplane  $H$ , and  $d_H$  is the linear coefficient associated with the hyperplane  $H$ . The objective function (16) aims to minimize the  $o$ -type norm of the sum difference between the approximate power generation of the hyperplane  $H$  and the exact value of the nonlinear HPF,  $\mathbf{PG}_k$ . Alternatively, for the bivariate case, we can rewrite the objective function as follows:

$$\min \Theta = \sum_{l=1}^{\mathbf{L}} \sum_{k=1}^{\mathbf{K}} \left| c^1_H \mathbf{Q}_k + c^2_H \mathbf{GH}_l + d_H - \mathbf{PG}_{kl} \right|^o \quad (17)$$

In (17),  $c^2_H$  is the slope or gradient 2 associated with  $\mathbf{GH}_l$  in the hyperplane  $H$ . After the fitting process,  $c^1_H$ ,  $c^2_H$ , and  $d_H$  will define the hyperplane  $H$  of the PWL. The next subsections deal with the MILP and MIQP optimization problems used to determine them. The first subsection deals with the models and the algorithm for obtaining nonconvex and concave univariate PWL. In the second subsection, the models for obtaining the concave PWL are expanded to the bivariate case. The models presented in this work require the numerical values of the domains of the variables to be properly limited so that the models present a good computational performance. Furthermore, big-M parameters are extensively used to model linear constraints. For reference, check [24].

Table 1 shows the summary of the models presented in this section. Checkmarks, “✓”, are for yes and “✗” are for no.

**Table 1.** Piecewise linear models.

	Piecewise Linear Model						
	1	2	3	4	5	6	7
Error Minimization	✓	✓	✗	✓	✗	✓	✗
Hyperplanes Minimization	✗	✗	✓	✗	✓	✗	✓
Univariate	✓	✓	✓	✓	✓	✗	✗
Bivariate	✗	✗	✗	✗	✗	✓	✓
Concave PWL	✗	✗	✗	✓	✓	✓	✓

3.1. PWL Models for the Univariate Case

3.1.1. PWL Model 1

The PWL model 1 seeks to determine a PWL function for the univariate plant-based HPF. Consider the optimization problem (18)–(36) reformulated from [24]:

$$\min \sum_{k=1}^K \left| c^1_H \mathbf{Q}_k + d_H - \mathbf{P}\mathbf{G}_k \right|^0 \tag{18}$$

$$\text{s.t. : } \mathbf{P}\mathbf{G}_k - (c^1_H \mathbf{Q}_k + d_H) \leq +\mathbf{M}_k^a (1 - \delta_{k,H}), \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \tag{19}$$

$$(c^1_H \mathbf{Q}_k + d_H) - \mathbf{P}\mathbf{G}_k \leq +\mathbf{M}_k^a (1 - \delta_{k,H}), \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \tag{20}$$

$$\sum_{H=1}^{\mathbf{B}-1} \delta_{k,H} = 1, \forall k = 1, \dots, \mathbf{K} \tag{21}$$

$$\delta_{k+1,H+1} \leq \delta_{k,H} + \delta_{k,H+1}, \forall k = 1, \dots, \mathbf{K}; H = 1, \dots, \mathbf{B} - 2 \tag{22}$$

$$\delta_{k+1,1} \leq \delta_{k,1}, \forall k = 1, \dots, \mathbf{K} - 1 \tag{23}$$

$$\delta_{k,B-1} \leq \delta_{k+1,B-1}, \forall k = 1, \dots, \mathbf{K} - 1 \tag{24}$$

$$\delta_{k,H} + \delta_{k+1,H+1} + \gamma_H - 2 \leq \delta_{k,H}^+, \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{25}$$

$$\delta_{k,H} + \delta_{k+1,H+1} + (1 - \gamma_H) - 2 \leq \delta_{k,H}^-, \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{26}$$

$$d_{H+1} - d_H \geq \mathbf{Q}_k (c^1_H - c^1_{H+1}) - \mathbf{M}_k^2 (1 - \delta_{k,H}^+), \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{27}$$

$$d_{H+1} - d_H \leq \mathbf{Q}_{k+1} (c^1_H - c^1_{H+1}) - \mathbf{M}_{k+1}^2 (1 - \delta_{k,H}^+), \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{28}$$

$$d_{H+1} - d_H \leq \mathbf{Q}_k (c^1_H - c^1_{H+1}) - \mathbf{M}_k^2 (1 - \delta_{k,H}^-), \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{29}$$

$$d_{H+1} - d_H \geq \mathbf{Q}_{k+1} (c^1_H - c^1_{H+1}) - \mathbf{M}_{k+1}^2 (1 - \delta_{k,H}^-), \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{30}$$

$$d_1 = 0 \tag{31}$$

$$c^1_H \in [\mathbf{C}^1_H, \overline{\mathbf{C}}^1_H], \forall H = 1, \dots, \mathbf{B} - 1 \tag{32}$$

$$d_H \in [\mathbf{D}_H, \overline{\mathbf{D}}_H], \forall H = 1, \dots, \mathbf{B} - 1 \tag{33}$$

$$\delta_{k,H} \in \{0, 1\}, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \tag{34}$$

$$y_H \in \{0, 1\}, \forall H = 1, \dots, \mathbf{B} - 2 \tag{35}$$

$$\delta_{k,H}^+, \delta_{k,H}^- \in [0, 1], \forall k = 1, \dots, \mathbf{K} - 1; H = 1, \dots, \mathbf{B} - 2 \tag{36}$$

In the optimization problem, **B** is the maximum number of breakpoints which is the first and last point of a line segment. The dimension of a hyperplane  $\in \mathbb{R}^2$  (our univariate

case) has dimension 1 and therefore is a line segment  $\mathbf{M}_k^a$  and  $\mathbf{M}_k^2$  are the non-negative big-M parameters,  $\delta_{k,H}$  is the binary variable that identifies which segment  $H$  each point  $k$  is associated with,  $\mathbf{C}_H^1$  and  $C_H^1$  are the minimum and maximum values of  $c^1_H$ ,  $\mathbf{D}_H$  and  $D_H$  are the minimum and maximum values of  $d_H$ ,  $\delta^{+/-}_{k,H}$  are continuous variables that indicate whether there is a breakpoint between  $\mathbf{Q}_k$  e  $\mathbf{Q}_{k+1}$ , and  $y_H$  is a binary variable that indicates the gradient change between adjacent linear segments.

The objective function (18) minimizes a chosen metric that, in this case, consists of the absolute error ( $\sigma = 1$ ) or the square error ( $\sigma = 2$ ); note that in the latter case, the model becomes a MIQP, which can considerably decrease the computational performance. The constraints (19) and (20) evaluate the approximation value at  $\mathbf{Q}_k$ . The constraint (21) ensures that each point  $k$  belongs to exactly one segment. The constraints (22)–(24) ensure the ordering of points in such a way that a point must be associated with the same or the next segment, where (23) and (24) ensures the ordering of the first and last data points, and (22) ensures the ordering of the other points. The PWL continuity is enforced through the constraints (25)–(30), which ensure that the breakpoint location,  $r_H$ , is equal in adjacent segments regardless of the gradient change. In short, if  $c_H - c_{H+1} > 0$ , it implies that  $y_H = 1$  and the constraints (27) and (28) are activated, which indicates that the gradient decreases between the two segments. Otherwise, if  $c_H - c_{H+1} < 0$ ,  $y_H = 0$ , the constraints (29) and (30) are activated, which implies that the gradient increases between the two segments. Finally, if  $c_H = c_{H+1}$ , there is no breakpoint at that location, and the function remains with the same gradient. The constraint (31) ensures at least one equation for which the plant power generation will be 0 whenever the plant turbined outflow is 0. Finally, the constraints (32)–(36) define the domain of variables.

The MILP optimal solution is given by the coefficients  $(c^1_H, d_H)/\forall H$ . However, since the resulting PWL is nonconvex, it is required that the location of each breakpoint,  $r_H$ , is known to delimit the domain for which each segment  $H$  is valid. These values are implicitly modeled in the optimization problem to ensure continuity and reduce the number of variables. The breakpoints' location can be calculated as:

$$r_1 = \mathbf{Q}_1 \tag{37}$$

$$r_H = \left\{ \begin{array}{l} \frac{d_{H+1}-d_H}{c^1_H-c^1_{H+1}}, \forall H = 2, \dots, \mathbf{B} - 1, c^1_H \neq c^1_{H+1} \\ \mathbf{Q}_k + \frac{\mathbf{Q}_{k+1}-\mathbf{Q}_k}{2}, \forall H = 2, \dots, \mathbf{B} - 1, c^1_H = c^1_{H+1} \end{array} \right\} \tag{38}$$

$$r_B = \mathbf{Q}_K \tag{39}$$

Above,  $r_1$  and  $r_B$  are the first and last breakpoint locations. For the other points, the location is calculated as in (38). Notice that if  $c^1_H = c^1_{H+1}$ , the model enforces  $d_H = d_{H+1}$ , which means that the two segments are affine with no discontinuity between  $(\mathbf{Q}_k$  e  $\mathbf{Q}_{k+1})$ ; thus,  $r_H$  can be chosen arbitrarily among these points.

### 3.1.2. PWL Model 2

The PWL model 2 has the same characteristics as Model 1 but with different strategies to establish the ordering and continuity of the function. Consider the optimization problem (40)–(56) reformulated from [25]:

$$\min \Theta = \sum_{k=1}^K |c^1_H \mathbf{Q}_k + d_H - \mathbf{P} \mathbf{G}_k|^{\sigma} \tag{40}$$

s.t.: (19)–(21), (31)–(34)

$$\delta_{k,H} = \delta_{k-1,H} + \delta^F_{k,H} - \delta^L_{k,H}, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \tag{41}$$

$$\sum_{k=1}^K \delta^F_{k,H} = 1, \forall H = 1, \dots, \mathbf{B} - 1 \tag{42}$$

$$\sum_{k=1}^{\mathbf{K}} \delta_{k,H}^L = 1, \forall H = 1, \dots, \mathbf{B} - 1 \quad (43)$$

$$\sum_{k'=1}^k \delta_{k',H}^F \geq \sum_{k'=1}^k \delta_{k',H+1}^F, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 2 \quad (44)$$

$$\sum_{k'=1}^k \delta_{k',H}^L \geq \sum_{k'=1}^k \delta_{k',H+1}^L, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 2 \quad (45)$$

$$\delta_{k,H}^F \leq \delta_{k,H}, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (46)$$

$$\delta_{k,H}^L \leq \delta_{k,H}, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (47)$$

$$\mathbf{Q}_k c_{H+1}^1 + d_{H+1} - (\mathbf{Q}_k c_H^1 + d_H) = p_{k,H}^+ - p_{k,H}^-, \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (48)$$

$$\mathbf{Q}_{k+1} c_H^1 + d_H - (\mathbf{Q}_{k+1} c_{H+1}^1 + d_{H+1}) = q_{k+1,H+1}^+ - q_{k+1,H+1}^-, \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (49)$$

$$p_{k,H}^+ \leq \mathbf{M}_k^a (1 - u_{k,H}), \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (50)$$

$$q_{k+1,H+1}^+ \leq \mathbf{M}_k^a (1 - u_{k,H}), \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (51)$$

$$p_{k,H}^- \leq \mathbf{M}_k^a (1 - v_{k,H}), \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (52)$$

$$q_{k+1,H+1}^- \leq \mathbf{M}_k^a (1 - v_{k,H}), \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (53)$$

$$u_{k,H} + v_{k,H} = \delta_{k,H}^L, \forall k = 1, \dots, \mathbf{K} - 1; \forall H = 1, \dots, \mathbf{B} - 2 \quad (54)$$

$$u_{k,H}, v_{k,H}, \delta_{k,H}^F, \delta_{k,H}^L \in \{0, 1\}, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (55)$$

$$p_{k,H}^{+/-}, q_{k,H}^{+/-} \in [0, \mathbf{M}_k^a], \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (56)$$

In this optimization problem,  $p_{k,H}^{+/-}$  and  $q_{k,H}^{+/-}$  are non-negative slack variables,  $u_{k,H}$  is a binary variable that takes on the value 1 if  $p_{k,H}^+ = q_{k+1,H+1}^+ = 0$ ,  $v_{k,H}$  is a binary variable that takes on the value 1 if  $p_{k,H}^- = q_{k+1,H+1}^- = 0$ ,  $\delta_{k,H}^F$  is a binary variable that assumes 1 if the point  $k$  is the first point in the segment  $H$ , and  $\delta_{k,H}^L$  is a binary variable that assumes 1 if point  $k$  is the last point of the segment  $H$ . Model 2 is identical to the model 1 concerning the objective function and constraints (19)–(21) and (31)–(34). The constraint (41)–(47) ensures the ordering, whereas (44)–(45) ensures the first and last points ordering of the data set. The ordering of the remaining points is handled via (46)–(47). The continuity of the HPF is enforced through the constraints (48)–(54). In this case, if a point is the last of the segment, (54) activates one of the binaries  $u_{k,H}$  or  $v_{k,H}$  that makes  $p_{k,H}^+$  and  $q_{k,H}^+ = 0$  or  $p_{k,H}^-$  e  $q_{k+1,H+1}^- = 0$ . Thus, this ensures continuity through (48) and (49) by ensuring that, at this point, both functions result in a non-negative or non-positive value. Finally, the constraints (55) and (56) define the variables domain.

### 3.1.3. PWL Model 3

The PWL Model 3 adapts the alternative formulation of [25] to minimize the number of breakpoints and, consequently, the number of segments for a given error.

$$\min \Theta = \sum_{H=1}^{\mathbf{B}-1} H \cdot W_H + 1000 \sum_{k=1}^{\mathbf{K}} F_k \quad (57)$$

s.t.: (21), (31)–(34), (41), (44)–(56)

$$\text{s.t. : } \mathbf{P}\mathbf{G}_k - (c_H^1 \mathbf{Q}_k + d_H) \leq \varepsilon_k + \mathbf{M}_k^a (1 - \delta_{k,H}), \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (58)$$

$$\left(c_H^1 \mathbf{Q}_k + d_H\right) - \mathbf{P}\mathbf{G}_k \leq \varepsilon_k + \mathbf{M}_k^a \left(1 - \delta_{k,H}\right), \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (59)$$

$$\varepsilon_k - F_k = \mathbf{E}r, \forall k = 1, \dots, \mathbf{K} \quad (60)$$

$$\sum_{k=1}^{\mathbf{K}} \delta_{k,H}^F - W_H = 0, \forall H = 1, \dots, \mathbf{B} - 1 \quad (61)$$

$$\sum_{k=1}^{\mathbf{K}} \delta_{k,H}^L - W_H = 0, \forall H = 1, \dots, \mathbf{B} - 1 \quad (62)$$

$$W_{H+1} - W_H \leq 0, \forall H = 1, \dots, \mathbf{B} - 1 \quad (63)$$

$$W_H \in [0, 1], \forall H = 1, \dots, \mathbf{B} \quad (64)$$

$$\varepsilon_k \in [0, \mathbf{M}_k^a], \forall k = 1, \dots, \mathbf{K} \quad (65)$$

$$F_k \geq 0, \forall k = 1, \dots, \mathbf{K} \quad (66)$$

In this optimization problem,  $W_H$  is a binary variable that is 1 if  $H \in \mathbf{B}$  is selected,  $F_k$  is a non-negative slack variable that only assumes a value greater than zero if the stipulated error  $\mathbf{E}r$  cannot be achieved, and  $\varepsilon_k$  is the absolute error at point  $k$ . The objective function (57) minimizes the number of breakpoints plus the sum of non-negative slack variables. In conjunction with the constraint (60), the problem aims to minimize the number of breakpoints for a given stipulated error  $\mathbf{E}r$  without using the slack variable ( $F_k = 0$ ), zeroing the second term of the objective function. If  $\mathbf{E}r$  cannot be achieved, then  $F_k > 0$ . The constraints (61) and (62) make  $\delta_{k,H}^F$  and  $\delta_{k,H}^L$  to be 0 for each inactive segment. The constraint (63) excludes some symmetric solutions. Lastly, constraints (64)–(66) delimit the domain of the introduced variables.

This formulation requires that a maximum value of breakpoints  $\mathbf{B}$  large enough is defined so that an optimal solution lies within the range from 2 to  $\mathbf{B}$  breakpoints.

There are cases where, depending on the region where the maximum error is observed, a maximum absolute error is required instead of having a low mean absolute error. If the error observed is in regions of interest of the HPF, it is possible to modify the constraint (60) and we can find a PWL as the slowest maximum absolute error for each point  $k$ .

$$\varepsilon_k - F_k = \mathbf{E}r \cdot \mathbf{P}\mathbf{G}_k, \forall k = 1, \dots, \mathbf{K} \quad (67)$$

With the support of the slack variable  $F_k$ , it is possible to establish an algorithm (Algorithm 1) with a searching heuristic to obtain the lowest possible number of breakpoints for the lowest error achieved.

---

**Algorithm 1.** The searching heuristic for the lowest number of breakpoints for a given error.

---

1. Choose the total number of points  $\mathbf{K}$ .
  2. Get the plant-based HPF data set  $(\mathbf{P}\mathbf{G}_k, \mathbf{Q}_k)$ .
  3. Choose the maximum number of breakpoints  $\mathbf{B}$ .
  4. Determine the metric-constraint (60) or (67).
  5. Choose an initial value of  $\mathbf{E}r$  and tolerance  $\text{tol}$ .
  6. While  $F_k \leq \text{tol}$ . Solve PWL Model 3.
  7. Do  $\mathbf{E}r = \mathbf{E}r - 0.01$  (for 1% reduction per run)
  8. Return the last set  $[c_H^1, d_H], \forall H = 1, \dots, \mathbf{B} - 1$  which  $\forall F_k \leq \text{tol}$
- 

### 3.1.4. PWL Model 4

This section discusses the generalizations necessary to obtain a concave PWL. Unlike the nonconvex PWL functions in which it is necessary to use binary variables to indicate the domain of each segment within the GS problem, a given value of a concave function according to [28] can be given by:

$$\mathbf{P}\mathbf{G}_k = f(\mathbf{Q}_k) = \text{minimum} \left\{ \mathbf{Q}_k c_H^1 + d_H \right\}, \forall H = \mathbf{B} - 1 \quad (68)$$

The objective function can be rewritten as:

$$\min \Theta = \sum_{k=1}^{\mathbf{K}} \left| \text{minimum} \left\{ \mathbf{Q}_k c_H^1 + d_H \right\} - \mathbf{P}\mathbf{G}_k \right|^0 \quad (69)$$

A well-known property of concave functions is that:

$$\mathbf{Q}_k c_H^1 + d_H \geq \mathbf{P}\mathbf{G}_k, \forall H = 1, \dots, \mathbf{B} - 1 \quad (70)$$

Constraints (70) are equivalent to removing the term  $\mathbf{M}_k^a(1 - \delta_{k,H})$  of (19).

$$c_H^1 - c_{H+1}^1 \geq 0, \forall H = 1, \dots, \mathbf{B} - 1 \quad (71)$$

The concave formulation has strong symmetry, which results in identical equations from the permutation of the coefficients  $c_H^1$  e  $d_H$ . The symmetry constraint (71) inhibits the existence of symmetric functions. However, it is worth noticing that the presence of constraint (71) might burden the model computationally so that it is advisable to treat the repeated coefficients externally. The presence of (70) makes the PWL models presented so far have solutions restricted to concave PWL. However, it makes several constraints in models 1, 2, and 3 unnecessary or redundant. Thus, the final formulation can be expressed as:

$$\min \Theta = \sum_{k=1}^{\mathbf{K}} \left| c_H^1 \mathbf{Q}_k + d_H - \mathbf{P}\mathbf{G}_k \right|^0 \quad (72)$$

s.t.: (20), (21), (31)–(34) and (71)

$$\mathbf{P}\mathbf{G}_k - \left( c_H^1 \mathbf{Q}_k + d_H \right) \leq 0, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (73)$$

Furthermore, relaxations of linear programming models that use big-M can have their performance improved if the following constraints are added to the model according to [28]:

$$c_H^1 \mathbf{Q}_k + d_H \geq 0, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (74)$$

$$\mathbf{P}\mathbf{G}_k \leq \sum_{H=1}^{\mathbf{B}-1} \left( c_H^1 \mathbf{Q}_k + d_H \right), \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (75)$$

### 3.1.5. PWL Model 5

This model restricts the PWL of Model 3 to concave functions.

$$\min \Theta = \sum_{H=1}^{\mathbf{B}-1} H \cdot W_H + 1000 \sum_{k=1}^{\mathbf{K}} F_k \quad (76)$$

s.t: (21), (31)–(34), (41), (44)–(47), (59)–(63), (71), (73)–(75)

$$\mathbf{P}\mathbf{G}_k - \left( c_H^1 \mathbf{Q}_k + d_H \right) \leq \varepsilon_k, \forall k = 1, \dots, \mathbf{K}; \forall H = 1, \dots, \mathbf{B} - 1 \quad (77)$$

The optimization problem above returns a concave PWL with the number of approximations less than  $\mathbf{B}$ , given the stipulated mean absolute error  $\mathbf{Er}$ . It is also possible to establish an algorithm (Algorithm 2) that seeks the lowest number of breakpoints for the lowest error reached, as carried out for the MLP 3.

---

**Algorithm 2.** The searching heuristic for the lowest number of breakpoints for a given error for the concave case.

---

1. Choose the total number of points  $\mathbf{K}$ .
  2. Get the plant-based HPF data set  $(\mathbf{P}\mathbf{G}_k, \mathbf{Q}_k)$ .
  3. Choose the maximum number of breakpoints  $\mathbf{B}$ .
  4. Determine the metric-constraint (60) or (67).
  5. Choose an initial value of  $\mathbf{E}r$  and tolerance  $\text{tol}$ .
  6. While  $Fk \leq \text{tol}$ . Solve PWL Model 5.
  7. Do  $\mathbf{E}r = \mathbf{E}r - 0.01$  (for 1% reduction per run)
  8. Return the last set  $[c^1_H, d_H]$ ,  $\forall H=1, \dots, \mathbf{B}-1$  which  $\forall Fk \leq \text{tol}$
- 

### 3.2. PWL Models for the Bivariate Case

#### 3.2.1. PWL Model 6

The PWL model 6 is intended to be used in the bivariate plant-based HPF, and it is an expansion of model 4. In the bivariate case, in addition to the coefficient,  $c^1_H$ , associated with turbined outflow  $\mathbf{Q}_k$ , it is necessary to calculate the coefficients  $c^2_H$  associated with the gross head  $\mathbf{G}\mathbf{H}_l$ . Note that, in the bivariate case, the HPF consists of hyperplanes  $H \in \mathbb{R}^3$  with its domain delimited by two variables, and therefore each approximation is the equation of a plane. The formulation to obtain the bivariate concave PWL is expressed as follows:

$$\min \Theta = \sum_{l=1}^L \sum_{k=1}^K \left| c^1_H \mathbf{Q}_k + c^2_H \mathbf{G}\mathbf{H}_l + d_H - \mathbf{P}\mathbf{G}_{kl} \right|^o \quad (78)$$

s.t: (31)–(33), (63), (64), (71)

$$\mathbf{P}\mathbf{G}_{kl} - (c^1_H \mathbf{Q}_k + c^2_H \mathbf{G}\mathbf{H}_l + d_H) \leq 0, \forall k = 1, \dots, \mathbf{K}; \forall l = 1, \dots, \mathbf{L}; \forall H = 1, \dots, \mathbf{N}\mathbf{H} \quad (79)$$

$$(c^1_H \mathbf{Q}_k + c^2_H \mathbf{G}\mathbf{H}_l + d_H) - \mathbf{P}\mathbf{G}_{kl} \leq \mathbf{M}_{kl}^a (1 - \delta_{kl,H}), \forall k = 1, \dots, \mathbf{K}; \forall l = 1, \dots, \mathbf{L}; \forall H = 1, \dots, \mathbf{N}\mathbf{H} \quad (80)$$

$$\sum_{H=1}^{\mathbf{N}\mathbf{H}} \delta_{kl,H} = 1, \forall k = 1, \dots, \mathbf{K}; \forall l = 1, \dots, \mathbf{L} \quad (81)$$

$$c_1^2 = 0, \quad (82)$$

$$c^1_H \mathbf{Q}_k + c^2_H \mathbf{G}\mathbf{H}_l + d_H \geq 0, \forall k = 1, \dots, \mathbf{K}; \forall l = 1, \dots, \mathbf{L}; \forall H = 1, \dots, \mathbf{N}\mathbf{H} \quad (83)$$

$$\mathbf{P}\mathbf{G}_{kl} \leq \sum_{H=1}^{\mathbf{N}\mathbf{H}} (c^1_H \mathbf{Q}_k + c^2_H \mathbf{G}\mathbf{H}_l + d_H), \forall k = 1, \dots, \mathbf{K}; \forall l = 1, \dots, \mathbf{L}; \forall H = 1, \dots, \mathbf{N}\mathbf{H} \quad (84)$$

$$c_H^2 \in [\mathbf{C}_H^2, \overline{\mathbf{C}}_H^2], \forall H = 1, \dots, \mathbf{N}\mathbf{H} \quad (85)$$

$$\delta_{kl,H} \in [0, 1], \forall k = 1, \dots, \mathbf{K}; \forall l = 1, \dots, \mathbf{L}; \forall H = 1, \dots, \mathbf{N}\mathbf{H} \quad (86)$$

In the optimization problem,  $\mathbf{N}\mathbf{H}$  is the maximum number of hyperplanes,  $\underline{\mathbf{C}}_H^2$  and  $\overline{\mathbf{C}}_H^2$  are the minimum and maximum values of  $c^2_H$ . The objective function (78) minimizes the chosen metric, the absolute error ( $o = 1$ ) or the square error ( $o = 2$ ). The constraints (79) and (80) evaluate the approximation value at the point  $(\mathbf{Q}_k, \mathbf{H}\mathbf{B}_l)$ . The constraint (79) also imposes concavity requirements. Constraint (81) ensures that each point belongs to only one plane. The constraints (82) add a second requirement (first, being (31)) to ensure that there is a plane only  $\mathbf{Q}$ -dependent so that the HPF returns a power generation 0 for any situation where  $Q = 0$ . The constraints (83) and (84) improve the model's computational performance. Finally, (85)–(86) delimit the domain of variables.

### 3.2.2. PWL Model 7

The PWL model 7 expands the PWL model 5 to the bivariate case. In this model, the objective is to minimize the number of planes for a certain pre-established error.

$$\min \Theta = \sum_{H=1}^{NH} H \cdot W_H + 1000 \sum_{l=1}^L \sum_{k=1}^K F_{kl} \tag{87}$$

s.t: (31)–(33), (71), (79)–(86)

$$\delta_{kl,H} = \delta_{kl-1,H} + \delta_{kl,H}^F - \delta_{kl,H}^L, \forall k = 1, \dots, K; \forall l = 1, \dots, L; \forall H = 1, \dots, NH \tag{88}$$

$$\sum_{l=1}^L \sum_{k=1}^K \delta_{kl,H}^F - W_H = 0, \forall H = 1, \dots, NH \tag{89}$$

$$\sum_{l=1}^L \sum_{k=1}^K \delta_{kl,H}^L - W_H = 0, \forall H = 1, \dots, NH \tag{90}$$

$$\sum_{l'=1}^l \sum_{k'=1}^k \delta_{kl',H}^F \geq \sum_{l'=1}^l \sum_{k'=1}^k \delta_{kl',H+1}^F, \forall k = 1, \dots, K; \forall H = 1, \dots, NH - 1 \tag{91}$$

$$\sum_{l'=1}^l \sum_{k'=1}^k \delta_{kl',H}^L \geq \sum_{l'=1}^l \sum_{k'=1}^k \delta_{kl',H+1}^L, \forall k = 1, \dots, K; \forall H = 1, \dots, NH - 1 \tag{92}$$

$$\delta_{kl,H}^F \leq \delta_{kl,H}, \forall k = 1, \dots, K; \forall H = 1, \dots, NH \tag{93}$$

$$\delta_{kl,H}^L \leq \delta_{kl,H}, \forall k = 1, \dots, K; \forall H = 1, \dots, NH \tag{94}$$

$$\varepsilon_{kl} - K_{kl} = Er, \forall k = 1, \dots, K; \forall l = 1, \dots, L \tag{95}$$

$$F_{kl} \geq 0, \forall k = 1, \dots, K; \forall l = 1, \dots, L \tag{96}$$

At last, it is also possible to establish a search algorithm (Algorithm 3) for the formulation (87)–(96).

---

**Algorithm 3.** The searching heuristic for the lowest number of breakpoints for a given error for the bivariate concave case.

---

1. Choose the total number of points **K** and **L**.
  2. Get the plant-based HPF data set (**PG<sub>kl</sub>**, **GH<sub>l</sub>**, **Q<sub>k</sub>**).
  3. Choose the maximum number of hyperplanes **NH**.
  4. Determine the metric-constraint (95) or (67).
  5. Choose an initial value of Er and tolerance tol.
  6. While  $F_{kl} \leq \text{tol}$ . Solve PWL Model 7.
  7. Do  $Er = Er - 0.01$  (for 1% reduction per run)
  8. Return the last set  $[c^1_H, c^2_H, d_H], \forall H=1, \dots, \mathbf{B}-1$  which  $F_{kl} \leq \text{tol}$
- 

### 3.3. Solution Quality Evaluation

After obtaining the PWL through the PWL model, it is necessary to measure the respective quality. The objective function of some of the proposed formulations returns the error, but only concerning the points used as a reference in obtaining the approximations. Thus, the PWL test for a larger data set is proposed to test the segments or planes obtained. This accuracy evaluation is performed by using the maximum absolute (97) and the mean absolute error (98).

$$\text{MAX\_A} = \max \left( \frac{|\mathbf{P}_n - \mathbf{PG}_n|}{|\mathbf{PG}_n|} \right), \forall n \in \mathbf{N} \tag{97}$$

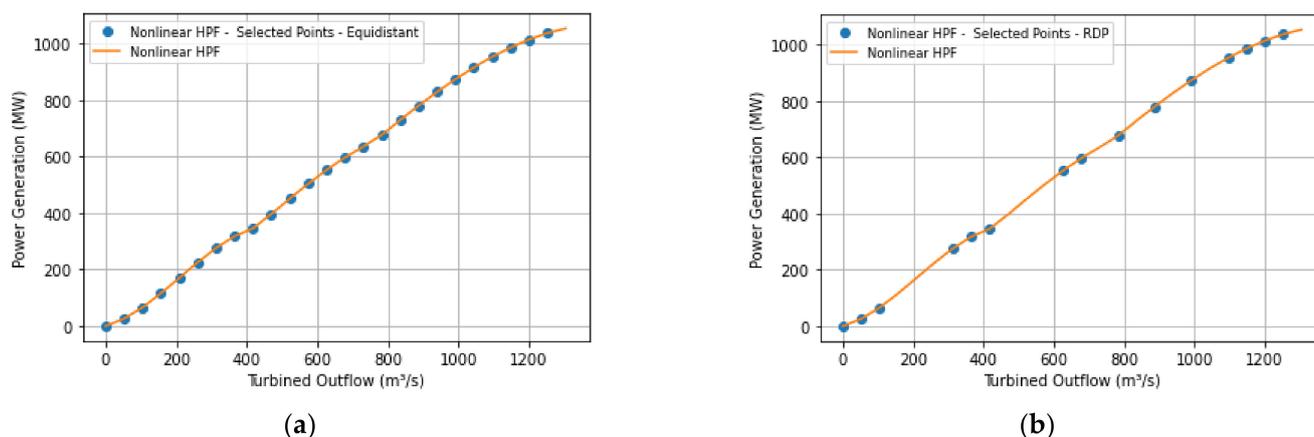
$$\text{MAE} = \frac{1}{\mathbf{N}} \sum_{n=1}^{\mathbf{N}} \frac{|\mathbf{P}_n - \mathbf{PG}_n|}{|\mathbf{PG}_n|} \tag{98}$$

In the optimization problem,  $n$  is the index for data point being tested,  $N$  is the number of data points,  $PG_n$  is the power generation at point  $n$ ,  $P_n$  is the power generation at point  $n$  obtained by employing the PWL, MAE is the mean absolute error, and MAX\_A is the maximum absolute error. Therefore, we employ a large number of  $N$  points for accuracy assessment.

#### 4. Optimized Data Selection through the Ramer-Douglas-Peucker Algorithm

The data obtained in Section 2 can be directly used in the PWL models presented in Section 3. The increase in the number of data and a maximum number of hyperplanes considerably increases the number of variables and constraints, bringing computational limitations to the model. In this aspect, using the Ramer-Douglas-Peucker algorithm [26,27] is suggested for optimized selection of points and consequent reduction of the number of data without excessively affecting the precision of the PWL obtained. The Ramer-Douglas-Peucker algorithm reduces the number of points on a curve approximated by several points. The algorithm eliminates points from a curve as long as a tolerance  $\epsilon$  is maintained to control this elimination. Thus, the plant-based HPF can reduce the number of points that constitute them. Note that if  $\epsilon$  is low enough, the algorithm will only eliminate unnecessary points from the curve, more precisely those that describe a region with strong linearity.

For example, consider the univariate HPF of a Hydropower plant (HP) with a discretization of  $K = 25$  values of  $Q_k$  and  $\epsilon = 2$  MW with the gross head  $GH$  fixed in 110 m. Figure 1 shows the selection of points with equidistant discretization without using the algorithm (a) and (b) the selection of points with the RDP algorithm.

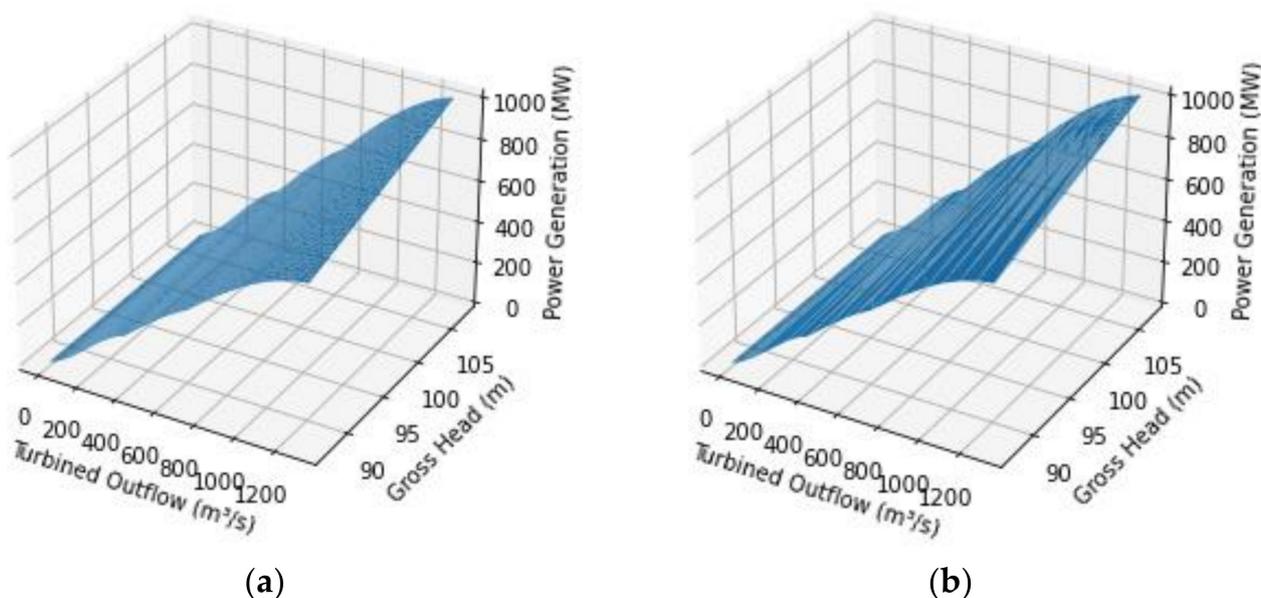


**Figure 1.** Data points selected for the univariate function: (a) without the RDP algorithm; (b) with the RDP algorithm.

The algorithm reduced the number of points to 15 in this example. Note that the exclusion of points occurs mainly in the more linear regions, thus not harming the accuracy of the PWL obtained.

A convenient way to use the algorithm is to choose a massive amount of data and let the algorithm itself select the points. Consider now the bivariate plant-based HPF of the same HP with a selection of 50 points of  $Q_k$  and 50 points in  $GH_l$ , which originates 2500 points  $PG_{kl}$ . Still, consider that  $\epsilon = 2$  MW. The application of the algorithm reduces the number of points to describe the curve from 2500 to 791 according to Figure 2.

Thus, RDP applications can be an important tool for data processing before using them in the proposed PWL models.



**Figure 2.** Data points selected for the bivariate: (a) without the RDP algorithm; (b) with the RDP algorithm.

## 5. Experimental Results

The models were implemented in python, embedded with GUROBI version 8.1.1. The codes were run on an Intel 2.5 GHz machine with 8 GB of RAM. The data sets were taken from two Hydropower Plant (HP) in Brazil: Santo Antônio Hydropower Plant (SAHP) and Machadinho Hydropower Plant (MAHP).

SAHP is the fifth-biggest hydropower plant in Brazil, with a total capacity of 3568 MW. The SAHP possesses 50 bulb turbines allocated into two groups: 24 four-blade with 73.29 MW and 26 five-blade with 69.59 MW, each group with a specific hydraulic efficiency curve. Table 2 describes the operating limits for the GUs.

**Table 2.** Operating limits for the GUs of SAHP.

	$w^{\min/\max}$	$GH^{\min/\max}$	$g^{\min/\max}$
Group 1	180/620	9/25.15	23.60/73.29
Group 2	180/568	9/20.69	22.40/69.59

On the other hand, the MAHP has three identical GUs with a 380 MW capacity. Table 3 shows the operating limits for the MAHP GUs.

**Table 3.** Operating limits for the GUs of MAHP.

	$w^{\min/\max}$	$GH^{\min/\max}$	$g^{\min/\max}$
Group 1	280/435	87.20/107.10	290/380

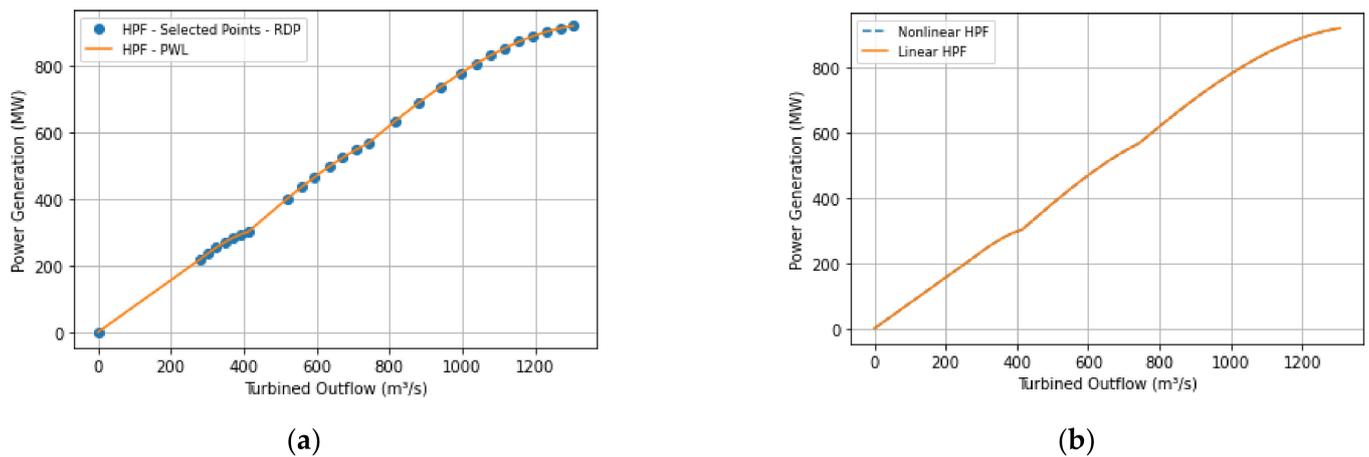
Section 5.1 analyzes the PWL models 1, 2, and 4 for fitting univariable discrete data with a predetermined number of breakpoints. In Section 5.2, the nonconvex and concave PWL functions are evaluated for the number of breakpoints minimization by employing the PWL models 3 and 5 (Algorithm 1 and Algorithm 2, respectively). Finally, in Section 5.3, we analyze the approaches intended for the bivariate case with a predefined number of planes.

### 5.1. PWL Models 1, 2 and 4

The optimization model (11)–(15) was first used to obtain  $K = 1000$  points  $(Q_k, PG_k)$  with  $GH = 100$  m for MHP and  $GH = 15$  m for SAHP. The RDP algorithm was then used

for optimal discretization. The resulting dataset was then applied to models 1, 2, and 4 to obtain the PWL. The number of breakpoints  $\mathbf{B}$  was fixed at 10, which resulted in nine segments. Finally, the evaluation was verified using the objective function value (sum of the absolute difference, where  $\sigma = 1$ ), the maximum absolute error, and mean absolute error (97) and (98) for  $\mathbf{N} = 1000$  points ( $\mathbf{Q}_n, \mathbf{P}\mathbf{G}_n$ ).

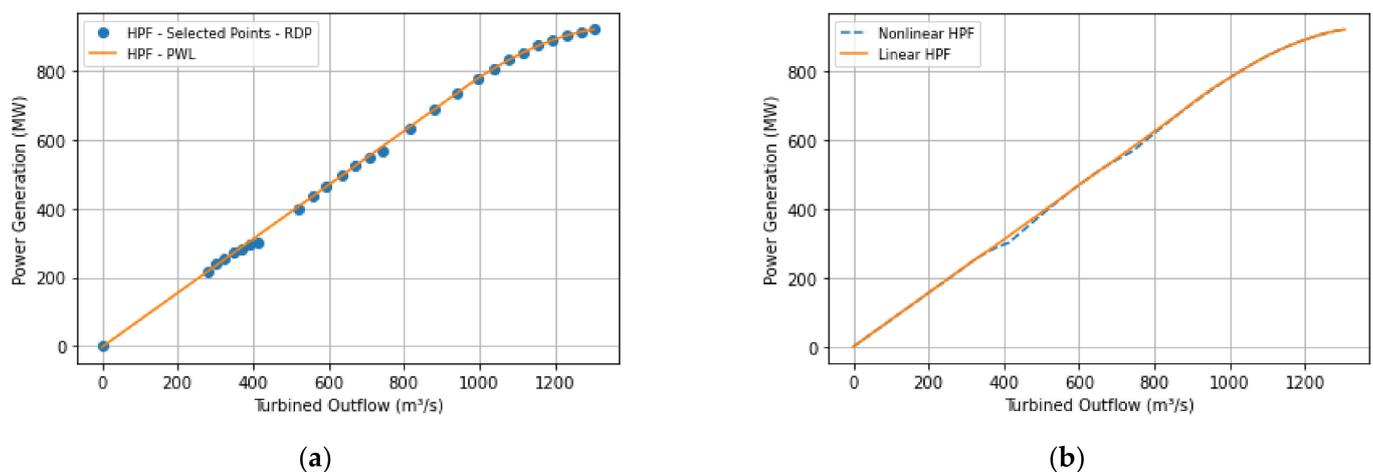
When applied to the 1000 initial points, the RDP algorithm selected 27 and 17 points ( $\mathbf{Q}_k, \mathbf{P}\mathbf{G}_k$ ), respectively, for MHP and SAHP with  $\varepsilon = 0.5$  MW. Figure 3a shows the points selected from the nonlinear HPF by the RDP algorithm and the behavior of model 1. Figure 3b illustrates the approximation behavior for the 1000 test points.



**Figure 3.** PWL function obtained by model 1 in MHP: (a) Selected points by the RDP algorithm and PWL approximation for the selected points. (b) Nonlinear HPF and Linear HPF obtained by using the PWL for 1000 test points.

Note that there is a region below  $280 \text{ m}^3/\text{s}$  in the nonlinear HPF in which no points were selected since it refers to the GU forbidden zone. In this case, a turbined outflow between 0 and  $280 \text{ m}^3/\text{s}$  cannot be supplied by the plant since even if the plant uses a single GU, this value would still be below the operating limits.

Figure 4 shows the curves obtained for the same data set for the PWL model 4 (concave approach).



**Figure 4.** PWL functions obtained by the PWL model 4 in MAHP. (a) Selected points by the RDP algorithm and PWL approximation for the selected points (b) HPF obtained for the PWL with 1000 test points.

Table 4 summarizes the three formulations applied to the MAHP.

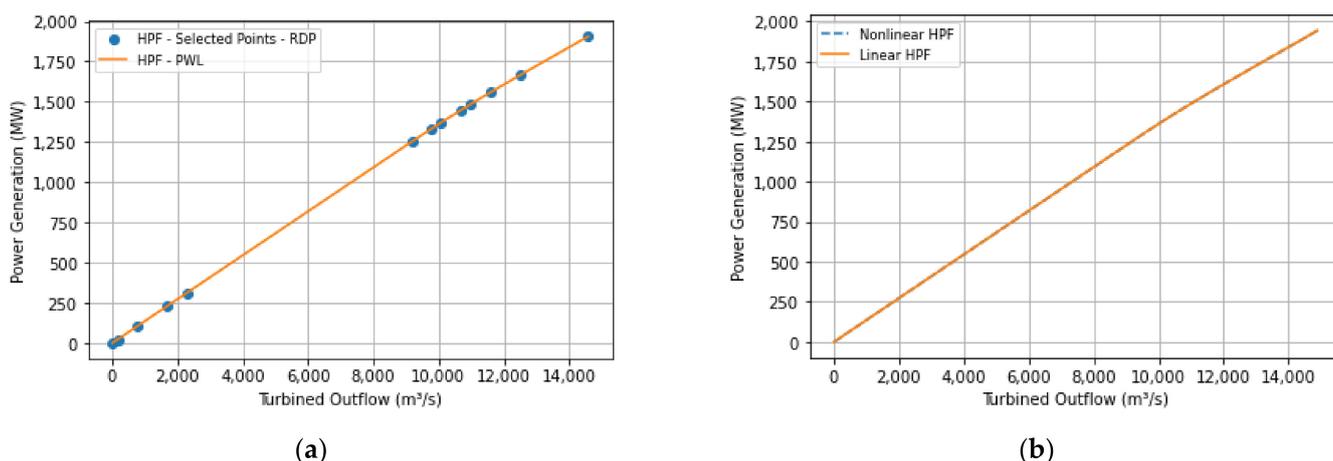
**Table 4.** PWL function evaluation for the MAHP.

DATA	Model 1	Model 2	Model 4
MAE (%)	0.064	0.063	0.742
MAX_A (%)	0.908	0.829	6.473
Running Time (s)	1.419	4.074	112.489

Regarding the models that generate nonconcave PWL functions, model 2 showed slightly better accuracy results. Model 1 proved to be superior in terms of time simulation time, which is expected, considering that model 2 presents a greater number of variables and constraints. However, it should be noted that if the linearization process is not embedded in the GS model, the simulation time factor may be negligible since it will be performed only once (offline). It is important to point out that the PWL model 2 did not always show better accuracy than model 1, and, therefore, it is advisable to run both models every time and choose the one that best fits the data.

It is worth mentioning that accuracy might improve for a different number of **B** and if the RDP tolerance  $\epsilon$  is narrowed. This action, however, would increase the running time of the model.

As a consequence of the concavity requirement, the PWL model 4 presented an error higher than the nonconvex models; however, it can be applied in GS models without binary variables. For group 2 of SAHP, model 2 yielded the behavior as seen in Figure 5.



**Figure 5.** PWL function obtained by model 2 in SAHP for the group 1 generating units. (a) Selected points by the RDP algorithm and PWL approximation for the selected points. (b) Nonlinear and linear HPF obtained by using the PWL for 1000 test points.

Tables 5 and 6 summarize the results for the three models applied to both groups of SAHP.

First, in general, the results were better for SAHP than MAHP. That is explained by the behavior of the nonlinear HPF of the latter. The more GUs a hydropower plant has, the more linear the plant-based HPF is expected to be, which directly impacts both precision and performance. The great linearity also explains the large exclusion of points in the intermediate (more linear) region.

**Table 5.** PWL evaluation for the SAHP Group 1.

DATA	Model 1	Model 2	Model 4
MAE (%)	0.043	0.040	0.098
MAX_A (%)	0.986	1.199	10.265
Running Time (s)	21.505	24.908	1.353

**Table 6.** PWL evaluation for the SAHP Group 2.

DATA	Model 1	Model 2	Model 4
MAE (%)	0.040	0.045	0.094
MAX_A (%)	1.336	0.822	8.855
Running Time (s)	14.902	17.193	1.258

The running time of model 2 was higher than model 1, again, explained by the high number of constraints and variables. In terms of accuracy, model 1 and 2 presented a very close mean absolute error for both groups, with model 2 presenting a slightly better accuracy for the mean and maximum absolute error in group 1 of SAHP. Once more, the variance in terms of accuracy justifies the use of both models.

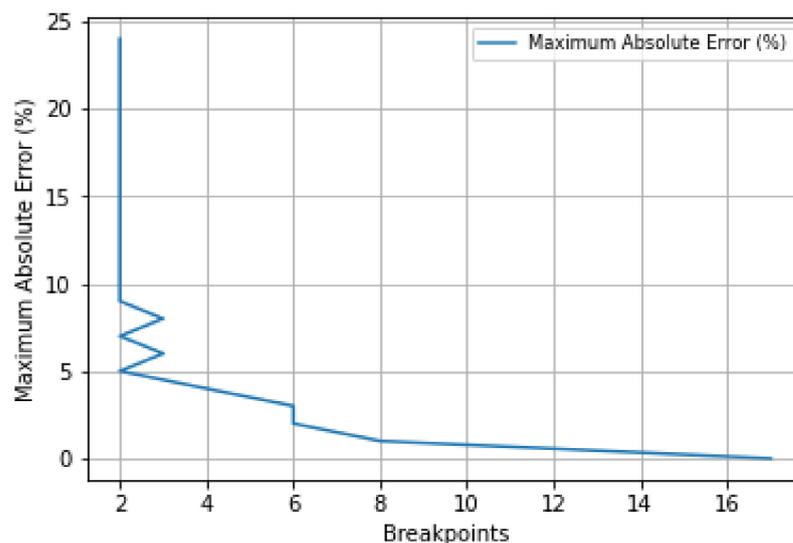
Both groups of SAHP showed a better mean absolute error than the MAHP but with a worse maximum absolute error. This performance is explained by the lower operational limit of the generating units of SAHP, where any difference in lower values of the curve incurs a bigger maximum absolute error. It is also worth noticing that the objective function minimized the mean absolute error for the given examples.

Despite reducing the number of variables and constraints concerning nonconvex models, we noticed in some simulations that the computational performance could sometimes be worse in concave models. This aspect is because the concavity condition constrains the breakpoints' location. On the other hand, the accuracy will always be equal (if nonlinear HPF is concave) or lower than those obtained in nonconvex models, mainly due to the error increasing in the nonconcave regions of the HPF.

### 5.2. PWL Models 3 and 5

For this type of analysis, the optimization model (11)–(15) is first used to obtain  $\mathbf{K} = 1000$  points ( $\mathbf{Q}_k, \mathbf{P}\mathbf{G}_k$ ) with the  $\text{GH} = 100$  m for MAHP and 15 m for SAHP. The RDP algorithm was then used for optimal discretization. The maximum number of breakpoints was stipulated as the maximum number of points obtained by the RDP (27 in MAHP and 17 in SAHP). Additionally, we considered  $\text{Er} = 0.25$  for the initial absolute error. Finally, Algorithm 1 and Algorithm 2 were applied to obtain the minimum breakpoints.

Figure 6 illustrates the progression of Algorithm 1 to reduce the maximum absolute error, using model 3 for MAHP. As can be seen, the error started at 25% with two breakpoints, reaching the optimal maximum absolute error at 0.7% with 17 breakpoints.



**Figure 6.** Algorithm 1—PWL model 3. Finding the minimum number of breakpoints for a given error for MAHP.

Table 7 summarizes the results obtained for both models.

**Table 7.** PWL functions evaluation for Algorithms 1 and 2 for MAHP.

DATA	Algorithm 1	Algorithm 2
B	17	5
MAE (%)	0.067	0.979
MAX_A (%)	0.700	7.352
Running Time (s)	195.117	2.389

Since we chose the maximum absolute error as the metric, it was observed that at the optimal number of 17 breakpoints, the maximum absolute error was better than at the fixed value of 10 breakpoints in Section 5.1. Tables 8 and 9 summarize the results for Algorithm 1 for SAHP.

**Table 8.** PWL function evaluation for Algorithms 1 and 2 for SAHP G1.

DATA	Algorithm 1	Algorithm 2
B	19	4
MAE (%)	0.349	0.077
MAX_A (%)	0.341	2.589
Running Time (s)	88.092	2.534

**Table 9.** PWL function evaluation for algorithm 1 and 2 for SAHP G2.

DATA	Algorithm 1	Algorithm 2
B	19	4
MAE (%)	0.401	0.042
MAX_A (%)	0.306	2.410
Running Time (s)	168.702	2.355

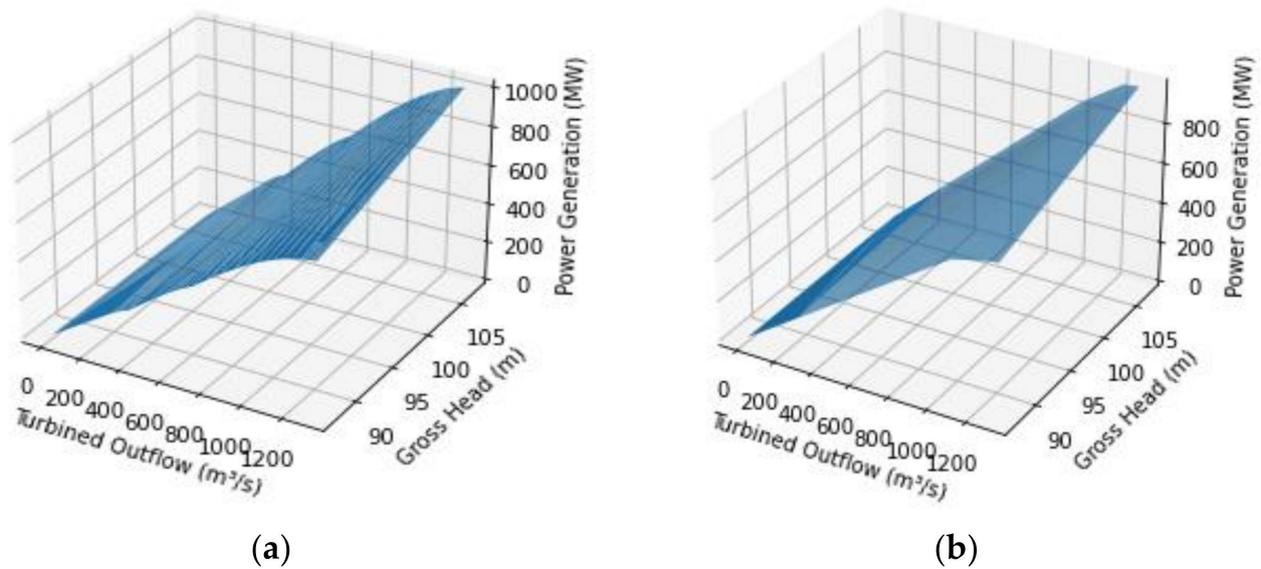
Algorithm 1, using PWL model 3, found an optimal solution of 0.341% and 0.306% for SAHP groups 1 and 2, respectively, with 19 breakpoints in both. Like MAHP, the mean absolute error was penalized when using the maximum absolute error as the stipulated error. Algorithm 2, using PWL model 5, found 2.589% and 2.410% maximum absolute error with four breakpoints.

### 5.3. Analysis of Models 6 and 7

The PWL models 6 and 7 are proposed to the bivariate case. The optimization model (6)–(10) was first used to obtain the dataset, with  $K = 100$  and  $L = 100$ , which resulted in 10,000 points  $(Q_k, GH_k, PG_k)$ . The RDP algorithm was used for optimal discretization of the points. The number of NH planes was set at 10 for formulation 6. Finally, the evaluation was performed similarly to Sections 5.1 and 5.2, but the PWL was tested for 10,000 data points.

For formulation 7, the maximum number of planes was stipulated as the maximum number of points obtained by the RDP and the initial error  $Er = 0.25$  (25%) for the maximum absolute error. Then, algorithm 3 was applied to obtain the minimum number of planes.

Applying the PWL model 6 for MAHP, we observed the following curve in Figure 7. Table 10 summarizes the results obtained with model 6.



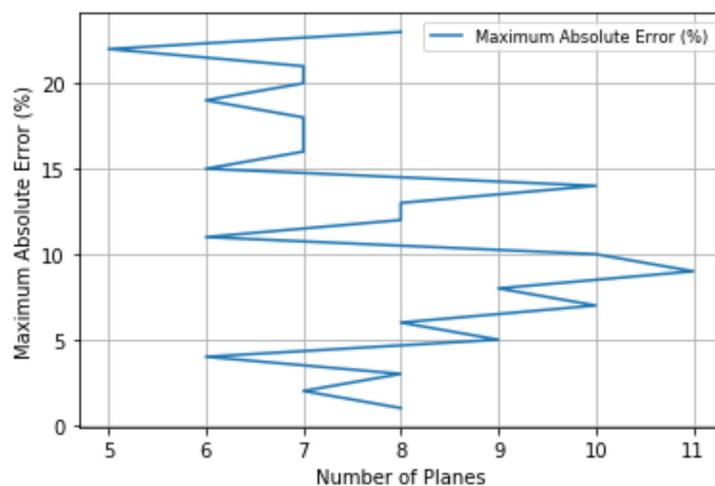
**Figure 7.** PWL functions for MAHP. (a) Planes obtained through the RDP selected points. (b) HPF obtained for the PWL with 10,000 test points.

**Table 10.** PWL model 6 results.

DATA	PWL Model 6		
	MAHP	SAHP G1	SAHP G2
MAE (%)	0.888	0.521	0.127
MAX_A (%)	6.356	4.097	4.456
Running Time (s)	1.108	0.204	2.391

As seen, in general, with more variables it is usual to observe a worsening of the precision. However, the mean absolute error was reasonable even for the bivariate case. The model convergence improved considerably since adding one more gradient added some malleability when obtaining the approximations. Using Algorithm 3, we noticed the following progressions for group 1 of SAHP.

The performance of Algorithm 3 applied to SAHP is presented in Figure 8, which stopped at 2.970% maximum absolute error to the reference points. Additionally, it resulted in 4.680% when applied to the 10,000 test points.



**Figure 8.** Algorithm 3—PWL model 7. Finding the minimum number of planes for a given error for SAHP G2.

Table 11 summarizes the results obtained with model 6.

**Table 11.** PWL model 7 results.

DATA	PWL Model 7		
	MAHP	SAHP G1	SAHP G2
NH	12	31	8
MAE (%)	1.583	0.304	0.324
MAX_A (%)	5.225	4.474	4.677
Running Time (s)	12.321	52.365	59.052

It was observed that just as in the univariate case, the maximum absolute error was improved at the cost of worsening the average absolute error.

Even in the bivariate case, the computational time was good due to the freedom in the number of hyperplanes chosen. Note that despite the algorithm being an iterative process, the computational time to determine the optimal number of hyperplanes for a given error was relatively fast, considering that Model 7 has a short running time.

## 6. Conclusions

To sum up, this paper proposed a framework to obtain piecewise linear approximations for plant-based HPF supported by recent mathematical contributions in the literature. The keyword for the proposed framework is “controllability”; the framework allows the control of the number of hyperplanes and, consequently, the number of variables and constraints of the PWL model by using two possibilities: minimizing the error for a fixed number of hyperplanes or minimizing the number of hyperplanes for a given error. Firstly, using the Rammer-Douglas-Peucker (RDP) algorithm for the pre-selection of the dataset led to gains in model performance by reducing the number of variables and constraints. Regarding the univariate nonconvex models, the mean absolute error was below 0.1% with few breakpoints, having a massive data grid as a test base. On the other hand, the univariate concave models had a mean absolute error below 1% even in the bivariate cases. For a plant with a large number of GUs such as UHSA, the mean absolute error was below 0.4%, explained by the linearity of the plant-based HPF. The maximum absolute error was used as an input parameter in the models for minimizing the number of hyperplanes, less than 5% for most cases. By having the number of breakpoints/hyperplanes free, the models have a high computational performance, allowing iterative algorithms to determine the optimal number of breakpoints/hyperplanes with less than 4 min of simulation time. For future work, the linear HPF reflects the nonlinear HPF; thus, a better representation of the nonlinear HPF can provide more meaningful approximations. One way to improve the representation of nonlinear HPF is to represent the hydraulic efficiency better. Nonconvex models have their field of use and are more accurate despite a more restricted use. It would be interesting to expand the nonconvex model to the bivariate case. Therefore, studying ways to establish continuity in parametric models and modeling the constraints as linear constraints is necessary.

**Author Contributions:** Conceptualization, E.C.F. and D.L.d.S.A.; methodology, E.C.F. and D.L.d.S.A.; software, D.L.d.S.A.; formal analysis, E.C.F. and D.L.d.S.A.; investigation, D.L.d.S.A.; data curation, D.L.d.S.A.; writing—original draft preparation, D.L.d.S.A.; writing—review and editing, D.L.d.S.A. and E.C.F.; visualization, E.C.F. and D.L.d.S.A.; supervision, E.C.F.; project administration, E.C.F.; funding acquisition, E.C.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by CNPq (National Council for Scientific and Technological Development), the Coordination for the Improvement of Higher Education Personnel (CAPES), and Santo Antônio Energia, via a research project registered under the code PD-06683-0119 with ANEEL (National Electric Energy Agency).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to fact that they belong to companies in the Brazilian electricity sector and are part of the commercial strategies of agents in this environment.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the study's design, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

CH	Convex Hull
GS	Generation Scheduling
GU	Generating Unit
HP	Hydropower Plant
HPF	Hydroelectric Production Function
LACC	Logarithmic Aggregation Convex Combination
LP	Linear Programming
MAE	Mean Absolute Error
MAHP	Machadinho Hydropower Plant
MAX_A	Maximum Absolute Error
MILP	Mixed-Integer Linear Programming
MINLP	Mixed-Integer Nonlinear Programming
MIQP	Mixed-Integer Quadratic Programming
NLP	Nonlinear Programming
PWL	Piecewise Linear
RDP	Rammer-Douglas-Peucker
SAHP	Santo Antônio Hydropower Plant

## References

- Wood, A.J.; Wollenberg, B.F.; Sheblé, G.B. *Power Generation Operation and Control*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2014.
- Hjelmeland, M.N.; Helseth, A.; Korpås, M. Medium-Term Hydropower Scheduling with Variable Head under Inflow, Energy and Reserve Capacity Price Uncertainty. *Energies* **2019**, *12*, 189. [[CrossRef](#)]
- Kong, J.; Skjelbred, H.I.; Fosso, O.B. An overview on formulations and optimization methods for the unit-based short-term hydro scheduling problem. *Electr. Power Syst. Res.* **2020**, *178*, 106027. [[CrossRef](#)]
- Conceição, W.C.; Marcato, A.L.M.; Ramos, T.P.; Filho, J.A.P.; Brandi, R.B.S.; David, P.A.M.S.; da Silva Júnior, I.C. Hydrothermal systems operation planning using a discretization of energy interchange between subsystems. *Electr. Power Syst. Res.* **2016**, *132*, 67–77. [[CrossRef](#)]
- Kang, C.; Chen, C.; Wang, J. An Efficient Linearization Method for Long-Term Operation of Cascaded Hydropower Reservoirs. *Water Resour. Manag.* **2018**, *32*, 3391–3404. [[CrossRef](#)]
- Fredo, G.L.M.; Finardi, E.C.; de Matos, V.L. Assessing solution quality and computational performance in the long-term generation scheduling problem considering different hydro production function approaches. *Renew. Energy* **2019**, *131*, 45–54. [[CrossRef](#)]
- Lin, Y.; Qiao, Y.; Lu, Z. Long-term generation scheduling for renewable-dominant systems concerning limited energy supporting capability of hydrogeneration. *IET Gener. Transm. Distrib.* **2022**, *16*, 57–70. [[CrossRef](#)]
- Brandi, R.B.S.; Ramos, T.P.; David, P.A.M.-S.; Dias, B.H.; da Silva Junior, I.C.; Marcato, A.L.M. Maximizing hydro share in peak demand of power systems long-term operation planning. *Electr. Power Syst. Res.* **2016**, *141*, 264–271. [[CrossRef](#)]
- Helseth, A.; Fodstad, M.; Mo, B. Optimal Medium-Term Hydropower Scheduling Considering Energy and Reserve Capacity Markets. *IEEE Trans. Sustain. Energy* **2016**, *7*, 934–942. [[CrossRef](#)]
- Hjelmeland, M.N.; Helseth, A.; Korpås, M. Impact of Modelling Details on the Generation Function for a Norwegian Hydropower Producer. *J. Phys. Conf. Ser.* **2018**, *1042*, 012010. [[CrossRef](#)]
- Marcelino, C.G.; Camacho-Gómez, C.; Jiménez-Fernández, S.; Salcedo-Sanz, S. Optimal Generation Scheduling in Hydro-Power Plants with the Coral Reefs Optimization Algorithm. *Energies* **2021**, *14*, 2443. [[CrossRef](#)]
- Moreno, S.R.; Kaviski, E. Daily scheduling of small hydro power plants dispatch with modified particles SWARM optimization. *Pesqui. Oper.* **2015**, *35*, 25–37. [[CrossRef](#)]
- Hamann, A.; Hug, G.; Rosinski, S. Real-Time Optimization of the Mid-Columbia Hydropower System. *IEEE Trans. Power Syst.* **2017**, *32*, 157–165. [[CrossRef](#)]

14. Nazari-Heris, M.; Mohammadi-Ivatloo, B.; Gharehpetian, G.B. Short-term scheduling of hydro-based power plants considering application of heuristic algorithms: A comprehensive review. *Renew. Sustain. Energy Rev.* **2017**, *74*, 116–129. [[CrossRef](#)]
15. Das, S.; Bhattacharya, A.; Chakraborty, A.K. Fixed head short-term hydrothermal scheduling in presence of solar and wind power. *Energy Strategy Rev.* **2018**, *22*, 47–60. [[CrossRef](#)]
16. Gupta, A.; Kumar, A.; Khatod, D.K. Optimized scheduling of hydropower with increase in solar and wind installations. *Energy* **2019**, *183*, 716–732. [[CrossRef](#)]
17. Li, C.; Wang, W.; Chen, D. Multi-objective complementary scheduling of hydro-thermal-RE power system via a multi-objective hybrid grey wolf optimizer. *Energy* **2019**, *171*, 241–255. [[CrossRef](#)]
18. Kang, C.; Guo, M.; Wang, J. Short-Term Hydrothermal Scheduling Using a Two-Stage Linear Programming with Special Ordered Sets Method. *Water Resour. Manag.* **2017**, *31*, 3329–3341. [[CrossRef](#)]
19. Brito, B.H.; Finardi, E.C.; Takigawa, F.Y.K.; Pereira, A.I.; Gosmann, R.P.; Weiss, L.A.; Fernandes, A.; de Assis Morais, D.T.S. Exploring Symmetry in a Short-Term Hydro Scheduling Problem: The Case of the Santo Antônio Hydro Plant. *J. Water Resour. Plan. Manag.* **2022**, *148*, 05021026. [[CrossRef](#)]
20. Liao, S.; Zhang, Y.; Liu, B.; Liu, Z.; Fang, Z.; Li, S. Short-Term Peak-Shaving Operation of Head-Sensitive Cascaded Hydropower Plants Based on Spillage Adjustment. *Water* **2020**, *12*, 3438. [[CrossRef](#)]
21. Brito, B.H.; Finardi, E.C.; Takigawa, F.Y.K.; Nogueira, P.L.R.; Morais, D.T.S.A.; Fernandes, A. Domain Partition of the Hydro Production Function for Solving Efficiently the Short-Term Generation Scheduling Problem. *IEEE Access* **2021**, *9*, 152780–152791. [[CrossRef](#)]
22. Löschenbrand, M.; Korpås, M. Hydro Power Reservoir Aggregation via Genetic Algorithms. *Energies* **2017**, *10*, 2165. [[CrossRef](#)]
23. Thaeer Hammid, A.; Awad, O.I.; Sulaiman, M.H.; Gunasekaran, S.S.; Mostafa, S.A.; Manoj Kumar, N.; Khalaf, B.A.; Al-Jawhar, Y.A.; Abdulhasan, R.A. A Review of Optimization Algorithms in Solving Hydro Generation Scheduling Problems. *Energies* **2020**, *13*, 2787. [[CrossRef](#)]
24. Rebennack, S.; Krasko, V. Piecewise Linear Function Fitting via Mixed-Integer Linear Programming. *Inf. J. Comput.* **2020**, *32*, 507–530. [[CrossRef](#)]
25. Kong, L.; Maravelias, C.T. On the Derivation of Continuous Piecewise Linear Approximating Functions. *Inf. J. Comput.* **2020**, *32*, 531–546. [[CrossRef](#)]
26. Douglas, D.H.; Peucker, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartogr. Int. J. Geogr. Inf. Geovis.* **1973**, *10*, 112–122. [[CrossRef](#)]
27. Ramer, U. An iterative procedure for the polygonal approximation of plane curves. *Comput. Graph. Image Process.* **1972**, *1*, 244–256. [[CrossRef](#)]
28. Toriello, A.; Vielma, J.P. Fitting piecewise linear continuous functions. *Eur. J. Oper. Res.* **2012**, *219*, 86–95. [[CrossRef](#)]