A comparison of alternative models for solving a non-linear single plant Hydro Unit Commitment problem

Alexandre Heintzmann[a,b], Christian Artigues[b], Pascale Bendotti[a], Sandra Ulrich Ngueveu[b], Cécile Rottner[a]

[a] *EDF Lab Paris-Saclay, 7 Bd. Gaspard Monge, 91120 Palaiseau, France*
[b] *LAAS-CNRS, Université de Toulouse, CNRS, INP, Toulouse, France*

Abstract

A wide range of real world optimization problems involves continuous decisions and non-linearities. Each non-linear component of such problems can be modeled either linearly or non-linearly, considering or not additional integer variables. This results into different modeling choices that can drastically impact the solution time and quality. In this paper, we evaluate representative modeling alternatives, including common models from the literature as well as new models featuring less common functions. The single plant Hydro Unit Commitment problem (1-HUC) is the considered non-linear use case. Among the non-linearities of the 1-HUC, we focus on those involved in the power production, more precisely the head effect and the turbine efficiency. The power is defined as a two-dimensional non-convex and non-concave function of the water flow and head decision variables, the latter being itself a one-dimensional concave function of the turbined volume. We consider both the general problem and a common special case, assuming that the water head is fixed. Several available solvers are used for each non-linear model and the best virtual solver is retained to focus on the model capabilities rather than on the solver performance. Based on the numerical experiments, three models stand out as the most efficient in terms of computational time, solution quality and feasibility, sometimes in a counter-intuitive manner. For each of these models, a solver is highlighted as the most adequate.

Keywords: Non-linear programming, non-linear modeling, Hydro Unit Commitment

1. Introduction

In the real world, systems involving continuous decisions and non-linearities are frequent. In the literature, optimizing such systems via mathematical programming using directly off-the-shelf solvers requires to choose between two main modeling alternatives of each non-linear component, namely either a linear or a non-linear model, yielding possibly additional integer variables. A

non-linear model usually represents more closely a physical system than a linear one, but tends to require a longer computing time, in particular when no convexity property applies. Within these two main modeling possibilities, there are still a lot of modeling choices to make. For illustration purposes, **Figure 1** shows a real-world continuous non-linear function on interval $[0; 10]$ and four alternative functions, amongst many others, to model it. Figure **(1a)** shows the real-world continuous non-linear function. We consider a situation where this function, while closely reflecting the reality, cannot be handled by the solvers, or at least not efficiently. Consequently, we need another modeling alternative, which can be one of the following. Alternative **(1b)** relies on a single non-linear function, simpler than the real-word one. Alternative **(1c)** uses a family of elementary non-linear functions. Alternative **(1d)** is based on a piecewise-linear (PWL) function. Alternative **(1e)** considers a finite set of discrete points. These alternatives have several differences, such as the type of function(s) involved: non-linear non-convex non-concave for **(1b)**, concave for **(1c)** and linear for **(1d)**, or the need of additional binary variables for **(1c)**, **(1d)** and **(1e)**. Thus, two alternatives can have different properties, meaning their solutions may be nothing alike. Facing real-world non-linear problems, it is crucial to chose the best modeling alternative.

The non-linear use case considered in this paper is the single plant Hydro Unit Commitment (1-HUC). For the 1-HUC, we consider a valley with a plant located between an upstream and a downstream reservoir. Water flows from the upstrem reservoir to the downstream reservoir, going through the plant which generates power. The aim is to schedule the power production of the plant in order to maximize the value of the valley. The non-linearities of the 1-HUC studied in this paper are the power as a non-linear function of the flow and the head, the latter being a non-linear function of the volume in the reservoirs. Such a problem appears to be a core element in a valley. First there exist real-world valleys restricted to a single plant. Second a valley with cascaded plants can be decomposed with respect to single plants. In this work, we focus on the modeling of non-linearities related to the economic value; we will not consider various other non-linearities even though they are in practice useful from an operational point-of-view. The idea is to study a simplified problem capturing sensitive non-linear aspects of the power production. To that end, we analyze the effect of different modeling alternatives using several performance indicators given hereafter. Incorporating several nonlinear functions would make this analysis impracticable. More precisely, the considered 1-HUC features two non-linearities: a one dimensional concave non-linearity and a two dimensional non-convex non-concave non-linearity. Such non-linearities can lead to various issues, which can be highlighted using different indicators, such as precision, feasibility, computing time or solution quality. The modeling of the hydroelectric power function already raised interest in the literature, e.g., in [16].

In this paper, we push further the study of the impact of non-linear modeling. To do so, we first propose a non-linear model which closely corresponds to a set of real world data. As this model is out of reach for current solvers, other models must be considered. Many different models are proposed in the litera-

ture. However, none of them feature a power function which can be considered as the exact analytic function corresponding to the physics. Hence, we propose seven different models. These models cover a large panel, ranging from linear to non-linear models, with and without integer variables, and are representative of how the literature handles the approximation functions described in **Figure 1**. More precisely, four models are proposed with a unique function as represented by **Figure (1b)**, and one is proposed for each of the alternatives represented in **Figure(1c)**, **(1d)** and **(1e)**. As any two non-linear solvers do not implement the same tools, they are expected to behave differently. This is why the proposed non-linear models will be solved using five black-box global optimization solvers available on Neos Server [12]. The principle is to evaluate these models for the 1-HUC using the indicators as defined in the previous paragraph. We then identify the non-linear features of the 1-HUC instances that impact their solution time and quality. The main contribution of the paper is to make general modeling recommendations based on the numerical experiments, depending on the features of the instance, the desired precision and the allowable computing time.

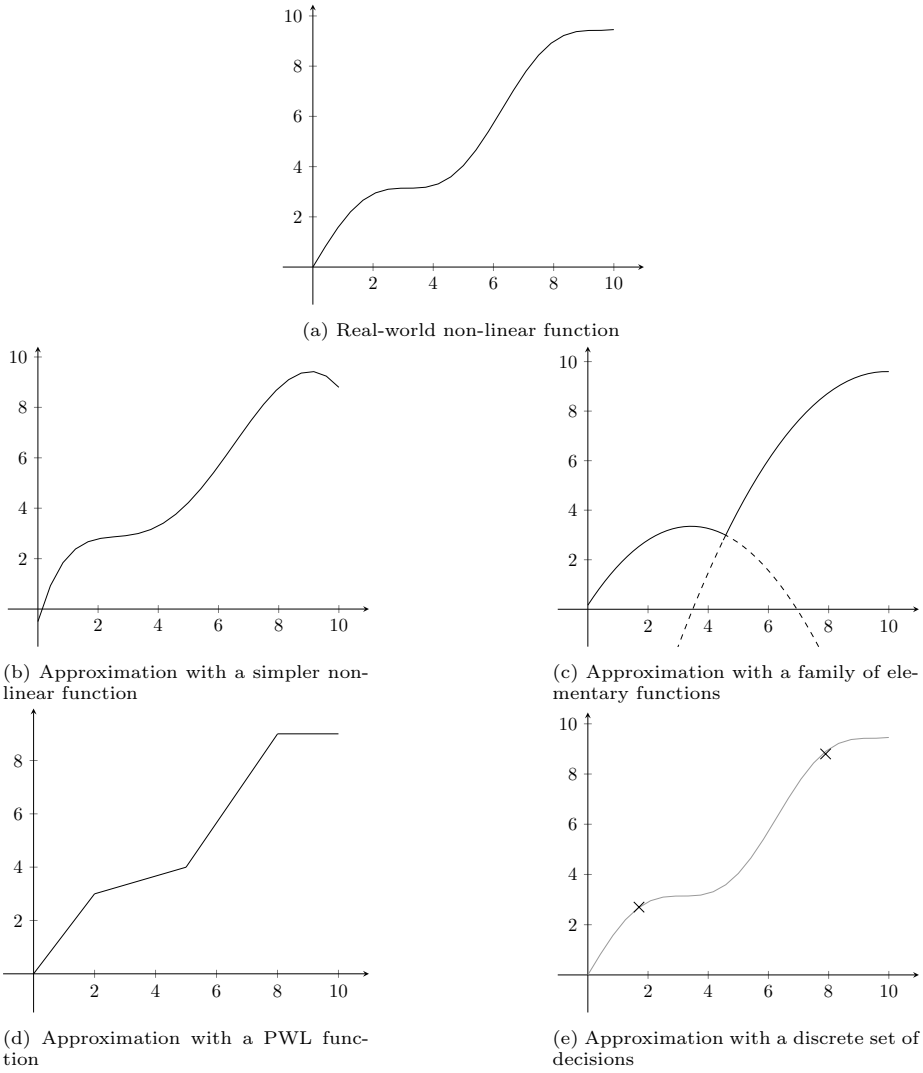
In Section 2 a literature review of solution approaches for non-linear optimization problems is proposed. In Section 3, the 1-HUC is defined, followed by a literature review on the non-linearities of the HUC. In Section 4 the proposed models are described and compared from a theoretical point of view. In Section 5, numerical experiments illustrate the comparative performance of the models on different sets of realistic 1-HUC instances. In Section 6, concluding remarks and perspectives for further research are drawn.

2. State-of-the-art and problem statement

In this paper, the focus is to evaluate different modeling alternatives for the 1-HUC solved directly by off-the-shelf exact solvers. Some solution approaches for non-linear problems and for the Hydro Unit Commitment will not be covered. In particular, we do not cover approaches such as instance decomposition [25], heuristics [38] or dynamic programming [4]. Hence, in this section, we first review the main generic modeling techniques, along with state-of-the-art solution methods implemented in the non-linear solvers used for the numerical experiments. Then, we review modeling alternatives described in the literature of hydro power systems and we present some recent modeling comparisons.

2.1. Solution approaches for (non)-linear optimization problems

As aforementioned, there are two main modeling possibilities to optimize a non-linear system using mathematical programming: either using a linear model or a non-linear model. Linear models lead to Mixed Integer Linear Programs (MILP), and non-linear models lead to Non-Linear Programs (NLP) or Mixed Integer Non-Linear Programs (MINLP). For these three types of programs, exact algorithms are based on a divide and conquer strategy: the search space is divided into sub-spaces for which upper and lower bounds can be obtained.



(a) Real-world non-linear function

(b) Approximation with a simpler non-linear function

(c) Approximation with a family of elementary functions

(d) Approximation with a PWL function

(e) Approximation with a discrete set of decisions

Figure 1: Four different approximations of a real-world non-linear function

For a minimization problem, the upper bound is derived from any feasible solution, and the lower bound is obtained by solving a relaxation of the problem. For an MILP, the relaxed problem generally consists in ignoring the integrity constraints [28]. For an NLP and an MINLP, the relaxed problem is obtained by ignoring the integrity constraints and also by replacing non-linearities by convex under-estimators [48].

Pointing out the differences between the models' families is also important because the tools used to solve a problem depend on its structure. To solve an MILP, the well known Branch and Bound (BB) algorithm [28] and its derivatives

the Branch and Cut [36], the Branch and Price [45] and the Branch and Cut and Price [13] can be used. To solve a convex (MI)NLP, as the relaxation is similar to a MILP, a variant of the BB algorithm can be used. To solve a non-convex (MI)NLP, a global optimisation algorithm is required. The main algorithm involved in most global optimization tools is the spatial Branch and Bound (sBB) [48], designed to solve an MINLP. If modeling requires no integer variables, the model remains continuous (**Figure (1b)**) and results in an NLP. Variations of the sBB for NLP can be used, such as the α Branch and Bound, the Reduced Space Branch and Bound [17], or the Branch and Contract [53]. If modeling requires additional integer variables (**Figure (1c)**), the algorithms involved in the tool must be able to handle an MINLP. As aforementioned, the sBB can solve this type of models as well as other algorithms such as the Branch and Reduce [43], the sBB with interval analysis [50] or the GMIN and SMIN algorithms [2].

2.2. Literature of the power and head function of the HUC

The power functon considered in this paper is a non-linear function of the flow and the water head, with the water head being the vertical distance between the water level of the upstream and the downstream reservoir. For this purpose, we introduce the following variables for each time period t:

- p_t : the continuous variable representing the power produced;
- d_t : the continuous variable representing the water flow;
- v_t^n : the continuous variable representing the volume in reservoir $n \in \{1, 2\}$;
- h_t : the continuous variable representing the head.

We also introduce functions $\boldsymbol{F}$ and $\boldsymbol{f}$ two generic functions such that $p_t = \boldsymbol{F}(d_t, h_t)$ and $h_t = \boldsymbol{f}(v_t)$. The aim of this review is to identify which functions are considered for $\boldsymbol{F}$ and $\boldsymbol{f}$ in the literature.

In [16], it is mentioned that there are cases of the HUC where no perfect analytic representation of the hydroelectric power function is known. Nevertheless, in various papers of the literature [6, 16, 29, 32, 35, 37], the shape of the power function is described as non-convex and non-concave, mainly due to the turbine efficiency. The turbine efficiency is non-linear with respect to the water flow d_t even with fixed-head H. Similarly, the turbine efficiency is non-linear with respect to the head h_t even for a fixed water flow d_t. We can then specify function $\boldsymbol{F}(d_t, h_t)$, with $\boldsymbol{g}$ the turbines efficiency.

$$\boldsymbol{F}(d_t, h_t) = \rho \cdot G \cdot h_t \cdot \boldsymbol{g}(d_t, h_t) \tag{1}$$

The power function is a product of ρ the density of water, G the gravitational constant, h_t the head and $\boldsymbol{g}(d_t, h_t)$ the turbines efficiency. Note that the function $\boldsymbol{g}$ can represent the efficiency of each turbine individually, or the efficiency of all turbines combined, depending on the modeling choice. In the following, we

review multiple ways of modeling the power function, taking into account the turbines efficiency.

A simplified hydroelectric power function in [23] is (1) with $\boldsymbol{g}(d_t, h_t) = d_t$:

$$\boldsymbol{F}(d_t, h_t) = \rho \cdot G \cdot h_t \cdot d_t$$

As such, function $\boldsymbol{F}$ is a bilinear function of the flow d_t and the head h_t. Note that a bilinear function is non-convex and non-concave [40], however it becomes linear when the head h_t is fixed. Hence, such a function does not correspond to the shape of the power function as described in the previous paragraph. Nevertheless, some (MI)NLP models of the HUC described in the literature features bilinear functions or a similar function. This is because a bilinear function is a common non-linear function, well handled by non-linear solvers even for large-scale instances. In [32] the HUC considered has multiple cascaded plants, and is modeled as an NLP, with the power defined as a bilinear function depending on the water flow and the head. In [30], an algorithm called spatial Hydro Branch and Bound (sHBB) has been developed to solve to optimality the HUC with cascading plants. This algorithm is used to solve an MINLP, where the power is a bilinear function of the head and the water flow.

A common modeling approach is to rely on polynomial functions, which are well known in the literature of MINLP. In [35] a more sophisticated function than a bilinear function is provided, where the power is represented by:

$$\boldsymbol{F}(d_t, h_t) = C_1 \cdot (v_t^1)^2 + C_2 \cdot (d_t)^2 + C_3 \cdot v_t^1 \cdot d_t + C_4 \cdot v_t^1 + C_5 \cdot d_t + C_6$$

with C_1 to C_6 being constants. This function is a sum of a bilinear function and two polynomials of degree two, depending on the volume and the water flow respectively. The formulation is further enhanced in [15], [18], [19] and [20] where the power function $\boldsymbol{F}(d_t, h_t)$ from [35] depicts the turbine efficiency, rather than the whole power function. In these papers, the whole power function is more complex, and also takes into account other non-linear functions, which we describe later in this review. In particular, the power function in these formulations can feature polynomial functions with degrees higher than 2. Using similar ideas, in [37] an NLP is presented, where the power function is approximated by a family of polynomials of degree two. More recently [7], a data-driven study of the non-linearities of the HUC has been conducted. The efficiency function considered is a piecewise polynomial function, where polynomial functions are not necessarily of degree two. In the same category of function, in [1] the power function of a unit features a fraction between polynomial functions:

$$\boldsymbol{F}(d_t, h_t) = \frac{d_t + C_1 \cdot (v_t^1)^2 + C_2 \cdot v_t^1 + C_3}{C_4 \cdot v_t^1 + C_5}$$

with C_1 to C_5 being constants.

Another common modeling approach uses a grid to have a reference for the hydroelectric power function obtained by discretizing the water flow and the head [16]. An algorithm is described in [46] to obtain a set of such grids, each

of them representing the power function for a given number of active units. For each point on one of these grids, the value of the hydroelectric power function is computed with a dynamic programming algorithm based on a bilinear function. The resulting grids are overvaluations of the power function, meaning that they do not necessarily reflect its actual shape.

Another modeling approach is to approximate the power function with a piecewise linear function. In [9] and [39], authors introduce a family of univariate PWL functions to model the power depending on the water flow. Each PWL function of this family represents the power with respect to the water flow for a specific volume. The model proposed in [6] for the 1-HUC also expresses the power with a family of univariate PWL functions of the water flow, for specific volumes. Besides, it takes into account the maximum variation of the water flow between two consecutive time periods. An improvement of this model features the rectangle method [14]. The aim is to compute a better approximation when v_t^1 is between two of the specific volumes selected to compute the PWL functions. To do so, the method computes a projection of the power between the two surrounding specific volumes in order to rectify the approximation. There are also iterative methods using PWL functions. In [21] the HUC with cascading plants is considered. It is pointed out that if the head is fixed, then the power depends only on the water flow. Using a PWL function with two pieces, the procedure is to solve HUC with fixed head iteratively, while updating the head between each iteration until convergence. A variant of the standard PWL models is to consider a PWL function approximating the convex hull of the power function [47]. In such a case, there is a loss of precision in concave parts, but the benefit is that there is no need for any binary variable. Indeed, for any water flow, the corresponding piece of the PWL function is always the one leading to the best value for the objective function. Besides PWL formulations, hyperplanes formulations have also been developed [42]. The aim is to create a set of hyperplanes for each number of active units, in order to linearize the non-linear power function. More precisely, the hyperplanes are deduced from the most efficient point, and each set forms a concave over-estimator of the power function for a given number of active units. As a maximization problem is considered, these hyperplanes yield a convex optimization region. Defining multiple sets aims to produce a more precise approximation, based on the aforementioned grid approximation of the power function for each number of active units [46]. For a given number of active units, the linearization does not require any additional binary variable. However, binary variables are required to indicate the number of active units at each time period, thus resulting in a MILP. Using the hyperplanes is in practice quite similar to using a PWL function approximating the convex hull of the non-linear function, presented previously for [47].

Comparing modeling alternatives raised interest in recent literature. In [16] three models for the hydro power maintenance scheduling are compared. The three models involve respectively a formulation with hyperplanes [42], a PWL formulation from [14], and a five degree polynomial function. The grid approach [46] is considered as a baseline to compare the economic value of the solutions. The result of this comparison is that the hyperplanes formulation is overall the

best compromise between the size and complexity of the optimization problem and the deviation from the reference data. It becomes relevant to include more sophisticated models in this comparison, as they may lead to smaller approximation errors. Another comparison has been made between modeling alternatives [15]. The three models considered are a model with high degree polynomials, a standard piecewise linear model, and a piecewise linear model of the convex hull of the power function. The results show that considering a non-linear model yields the best trade-off, as it requires a lower computational time than the standard PWL model, but a higher objective value than the convex hull PWL model. Another comparison has been carried out in [18], but instead of comparing modeling alternatives, the following three solution approaches are compared: solving a non-linear HUC with a Lagrangian relaxation, solving a non-linear HUC with the AIMMS outer approximation algorithm, and solving a linearized HUC with a MILP solver. The results show that the first two options yield the best results in terms of objective functions, while solving the MILP model deteriorates the objective function by 1 to 2%. However, the computational times of solving the MINLP model are much higher than the other options.

When it comes to the head effect, few alternatives are considered in the models of the literature. Indeed, modeling the head effect involves polynomial functions as described for instance in [1, 15, 18, 19, 29, 32]. No other modeling alternative is proposed for the head effect.

2.3. Problem statement

As the focus is to represent the power function, the 1-HUC considered in this paper is simplified with respect to other components. Thus, many constraints from the literature, listed hereafter, will be ignored.

The spillage [8, 32, 39] is the process of discharging water from the upstream reservoir to the downstream reservoir without going through the plant. The spilled water does not play a major role in the economic value of the valley. Indeed, it directly goes to the downstream reservoir, without activating any turbine, hence no energy is produced by the spilled water. The main purpose of spilling water is to avoid overfloods. Ramping constraints [14] limit the variation of the water flow between two consecutive time periods. These constraints are used in practice in order to take into account several other uses of water in the valley. In some models, there are start-up and shut-down costs [21, 30], making the start up and shut down of a unit impact the profit. We also consider a fixed unit start-up sequence [16]. This makes it possible to compare models aggregating all units, but also models where each unit is represented individually. Without such a fixed sequence, comparing models would become more difficult because the efficiency of a turbine can be impacted by the sequence. Also, some plants can have units operating in reverse in order to pump the water from the downstream reservoir to the upstream reservoir but this case will not be considered in this paper.

As we focus on the power function, we also consider a simplified head function. In the following, we consider the gross head, which does not take into

account the penstock losses [7]. The penstock loss is the power loss due to the friction between the water and the penstock. We also ignore the tailrace level [7], which depends on the water flow. As such, the head considered solely depends on the volumes, which allows for a more comprehensive study of the power function with respect to the water flow. More specifically, in such a case, the head is the height difference between the surface of the water in the upstream reservoir and the downstream reservoir. However, in the 1-HUC there is no control on the downstream reservoir as there is no downstream unit. In this problem, we focus the study of the head effect on the sensitivity to changes in the height of the upstream reservoir, rather than on the difference between upstream and downstream reservoirs. Since there is no unit downstream to turbine the coming water, the effect could indeed be amplified. For a more comprehensive study, considering a downstream unit would be needed which is beyond the scope of this work.

In this paper, we compare modeling alternatives, such as in [15] and [16], but not solution techniques as done in [18]. To do so, we consider generic models corresponding to the three common alternatives to represent the power function in the literature, namely a PWL function, a bilinear function or polynomial functions. We also consider generic models with non-linear functions that have not yet been studied in the literature of the HUC. As such, we push further the comparison done in [15] and [16], by considering a wider variety of models. Indeed, LP, MILP, NLP and MINLP models are considered, where some represent the power function of the whole plant, while others represent the power function of each unit explicitly. Besides, all models considered will also be compared for a standard simplification of the 1-HUC where the head is a constant.

3. 1-Hydro Unit Commitment

3.1. Definition of the problem

The 1-HUC is defined as follows. Consider a valley containing a hydro power plant located between a single upstream reservoir and a single downstream reservoir. From the hydroelectric production principle, the water from the upstream reservoir flows through the plant to the downstream reservoir, operating the N units of the plant. A unit is a combination of a turbine and a generator. When operating the units, electric power is generated. The units are not necessarily identical, and have a prescribed start-up order. The time horizon is discretized in T time periods, each of duration Δ. At each time period t the water flow d_t going through the units must lie within the interval $[\underline{D}, \overline{D}]$. The power p_t produced at the time period t depends on the water flow d_t, but also on the reservoir head h_t. The head h_t is the height difference between the surface of the water in the upstream reservoir and the plant, and depends on the volume of the upstream reservoir v_t^1. Each reservoir $n \in \{1, 2\}$ has a maximum capacity $\overline{V}_t^n$ and minimum capacity $\underline{V}_t^n$, variable through time. Variable maximum and minimum capacity makes it possible to set target volumes for specific time periods, when $\overline{V}_t^n = \underline{V}_t^n$. We detail the purpose of target volumes later in this

section. At each time period, the reservoir n has an additional intake of water A_t^n, which can be positive (rain, melting snow) or negative (use of water for surrounding agriculture). At each time period, we consider the energy to be sold at forecasted unitary price Λ_t, variable through the time. At the end of time period T, the water in reservoir n has an expected unitary value Φ^n. Value Φ^n is the expected value of the energy produced using the water. A higher Φ^n will lead to preserve more water and produce less energy, and the other way around for a lower Φ^n.

At Electricité de France (EDF), the HUC is considered as a revenue-maximizing price-taker scheduler problem, where the power prices, the water expected value, the reservoir external inflows and the reservoirs capacities are parameters. This is because the Unit Commitment Problem, which schedules the national production in order to meet the demand, is solved at EDF by a Lagrangian decomposition [41]. This decomposition yields sub-problems of the same nature (thermal, hydraulic, solar, ...). The prices, Φ^n and Λ_t in the case of the HUC, are given by the master problem.

Introducing target volumes allows to take into account water management policies. At EDF, these target volumes are defined on a year ahead basis to make the best use of water resource from an economical perspective. In this case, target volumes exist for the last time period of each day. Hence, depending on the number of days considered, there are up to one target volume for each day considered.

There exist other types of HUC, for instance where the aim is for the energy produced to meet the demand [21]. The work of our paper extends to these variants of the HUC because the non-linear hydro production functions are the same.

The profit takes into account the total value of the water in each reservoir at the end of time period T, and the value of the energy produced. Solving the HUC consists in maximizing the profit, while satisfying the capacities and the target volumes at each time period.

A generic model (P_{gen}) can be defined, using the water flow d_t, the power p_t, the volume in the upstream reservoir v_t^1, the volume in the downstream reservoir v_t^2 and the head of the upstream reservoir h_t as decision variables. As introduced in **Section 2.3**, function $\boldsymbol{f}$ gives the water head h_t at time period t with respect to the volume and function $\boldsymbol{F}$ represents the power of the units, depending on the water flow d_t and the head h_t. Function $\boldsymbol{f}$ is a one-dimensional concave function and $\boldsymbol{F}$ is a two-dimensional non-convex and

non-concave function. Model (P_{gen}) is given as follows:

$$\max \sum_{t=1}^{T} \Delta \cdot \Lambda_t \cdot p_t + \sum_{n=1}^{2} \Phi^n \cdot v_T^n \tag{2}$$

$$\text{s.t. } v_t^1 = V_0^1 + \sum_{t'=1}^{t} \left(A_{t'}^1 - d_{t'} \Delta \right) \qquad \forall t \leq T \tag{3}$$

$$v_t^2 = V_0^2 + \sum_{t'=1}^{t} \left(A_{t'}^2 + d_{t'} \Delta \right) \qquad \forall t \leq T \tag{4}$$

$$h_t = \boldsymbol{f}(v_t^1) \qquad \forall t \leq T \tag{5}$$

$$p_t = \boldsymbol{F}(d_t, h_t) \qquad \forall t \leq T \tag{6}$$

$$\underline{V}_t^n \leq v_t^n \leq \overline{V}_t^n \qquad \forall t \leq T, \forall n \in \{1, 2\} \tag{7}$$

$$\underline{D} \leq d_t \leq \overline{D} \qquad \forall t \leq T \tag{8}$$

Constraints (3) and (4) are volume conservation constraints. Note that it is also possible to define these constraints with consecutive time periods, which yields the same relaxation and similar computational times. Constraints (5) express the water head h_t, using the concave function $\boldsymbol{f}$ of the volume v_t^1. Constraints (6) define the power p_t, using the non-convex non-concave non-linear function $\boldsymbol{F}$ of the water flow d_t and the head h_t. Constraints (7) and (8) give upper and lower bounds for variables. The criterion to maximize is the profit, which is a linear expression described by (2).

This model acts as a framework for the various models proposed in this paper, hence (P_{gen}) is as generic as possible, functions $\boldsymbol{f}$ and $\boldsymbol{F}$ being not specified. We can still be more specific about some characteristics of (P_{gen}). Clearly, the volume v_t^n is non-negative. The water flow d_t is also non-negative because we only consider units with turbines. By definition of the head and the power, both functions $\boldsymbol{f}$ and $\boldsymbol{F}$ are non-decreasing and non-negative. This means that variables h_t and p_t are also non-negative. Besides, we consider in this case a standard non-linearity for the turbines efficiency, which is a concave function for each unit.

A standard simplification of the 1-HUC is to assume a constant head $h_t = H$, which leads to the fixed-head 1-HUC. This simplification is relevant for some instances of the 1-HUC where volume variations are small enough for the impact on the turbines efficiency to be insignificant. In such a case, equality (5) and (6) from (P_{gen}) are replaced by:

$$p_t = \boldsymbol{F}(d_t, H) \qquad \forall t \leq T \tag{9}$$

therefore, function $\boldsymbol{F}$ becomes a one-dimensional function, but remains non-convex non-concave. Note that even if the head is constant, we still consider variables v_t^1 and v_t^2 in the model, to ensure that reservoir capacities are respected.

3.2. Non-linear reference model for the 1-HUC

In this section, we describe a reference model (P_{ref}) for the 1-HUC. This model will be considered in the remaining of this paper as the real-word model (see **Figure 1a**). To define (P_{ref}), we consider the generic model (P_{gen}), to which we specify functions $\boldsymbol{f}$ and $\boldsymbol{F}$.

Firstly, we will focus on function $\boldsymbol{f}$ to define the head. **Figure 2a** shows the evolution of the head depending on the volume for a realistic instance named B-T-1, described in **Appendix C**. More generally, function $\boldsymbol{f}$ has the following form:

$$h_t = \boldsymbol{f}(v_t^1) = \gamma_1 + \gamma_2 \cdot v_t^1 + \gamma_3 \cdot (v_t^1)^{\gamma_4} \qquad \forall t \leq T \tag{10}$$

where γ_i are parameters depending on the instance, and $\gamma_4 \in [0.5, 1]$, which means that this function is necessarily concave. Depending on the shape of the reservoir, the function can be quasi linear or have a very noticeable non-linearity, but always stays concave.

Secondly, we will focus on the power function $\boldsymbol{F}$, featuring the turbine efficiency $\boldsymbol{g}$ as in equation (1). There is no analytic function $\boldsymbol{g}$ that perfectly represents the physics. In the following we define function $\boldsymbol{g}$ as well as possible, based on data and on the following information about the shape of the hydroelectric power function.

When the head is fixed, the power is the product of function $\boldsymbol{g}$ and constants, but remains non-convex and non-concave. The reason why $\boldsymbol{g}$ is non-convex and non-concave is because the function $\boldsymbol{g}$ includes the turbine efficiency of each unit. In particular, the turbine efficiency is concave, as represented in [3], and the plant has N units which start-up in a prescribed order. Adding multiple units adds non-concavity to the resulting function. To push the analysis further, function $\boldsymbol{g}$ also has the following characteristics. For each unit, function $\boldsymbol{g}$ is almost linear when the unit starts, then it incurves more and more until the next unit starts. When another unit starts, we notice a break in the function shape. The four main characteristics of $\boldsymbol{g}$ are described in **Table 1**.

Table 1: Characteristics of the power function

C1	non-convex and non-concave
C2	locally linear when a unit starts
C3	concave for each unit with respect to the water flow
C4	non-differentiable points when starting up a new unit

In order to be as close to the physics as possible, we define $\boldsymbol{g}$ as a piecewise non-linear function with N different five parameter logistic functions ($\boldsymbol{5PL}$) [24]. $\boldsymbol{5PL}$ functions are described in **Appendix B**. The $5PL$ i represents the power produced by the first i units combined with respect to the waterflow d_t. In the general case, function $\boldsymbol{g}$ also depends on the head. In particular, function $\boldsymbol{g}$ for the maximum head is not a linear transformation of function $\boldsymbol{g}$ for the minimum head. First, the turbine efficiency increases non-linearly with respect

to the head. Second, with a larger head, the apex of the efficiency curve is obtained with a higher water flow. Consequently, with a larger head one can use a turbine on a wider range of water flows, meaning that even the water flow d_t to start-up the units is larger. To take these effects into account, the parameters of the $\mathbf{5PL}$ functions are dependent on the head h_t.

We define $\boldsymbol{g}$ as a piecewise non-linear function. To do so, we define the following notations:

- a_{it} : the binary variables such that $a_{it} = 1$ if we use the $5PL$ function associated with the first i units at time period t;
- x_{jit} : the continuous variables being the j^{th} parameter of the $5PL$ function for the first i units at time period t;
- α_{ji} and β_{ji} : constants such that x_{jit} linearly depends on h_t with these parameters.

With these notations, we introduce the following constraints:

$$x_{jit} = \alpha_{ji} + \beta_{ji} \cdot h_t \qquad \forall j \leq 5 \forall i \leq N, \forall t \leq T \tag{11}$$

$$\sum_{i=1}^{N} a_{it} \leq 1 \qquad \forall t \leq T \tag{12}$$

$$a_{it} \in \{0, 1\} \qquad \forall i \leq N, \forall t \leq T \tag{13}$$

$$p_t = \boldsymbol{F}(d_t, h_t) = \rho \cdot G \cdot h_t \cdot \sum_{i=1}^{N} a_{it} \cdot \mathbf{5PL}(d_t, x_{1it}, \ldots, x_{5it}) \qquad \forall t \leq T \tag{14}$$

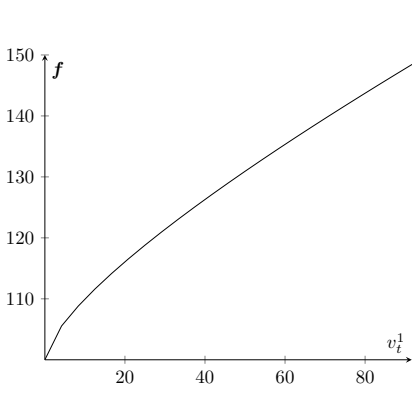
Equalities (11) define parameters x_{jit} to be linearly dependent on h_t. Constraints (12) and (13) ensure that at most one $5PL$ function is considered at each time period. Equalities (14) are the power at time period t with $\boldsymbol{g}$ as a piecewise non-linear function.

The complete MINLP model (P_{ref}) features objective function (2), constraints (3)-(4), (7)-(8) and (10)-(14). The $\mathbf{5PL}$ functions are parameterized such that only the concave part of the $\mathbf{5PL}$ is considered when i units are active, as well as locally linear when a unit starts. Besides, there is a non-differentiable point when starting up a new unit. As the power is clearly non-convex and non-concave, the non-linear power function in (P_{ref}) features the four main characteristics described in **Table 1**.

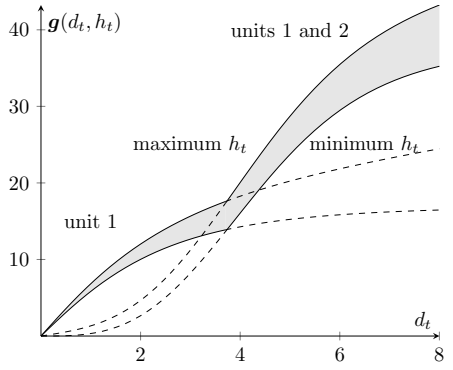
When considering a piecewise (non-)linear model, one usually requires additional constraints to match the values of the binary variables corresponding to the pieces with the values of the decision variables. For model (P_{ref}), this means adding constraints to indicate, for a given time period t, which binary variable a_{it} is equal to 1, depending on d_t. However, in some cases, there is no need to describe which binary variable must be equal to 1, and sometimes binary variables are not even required such as for convex PWL functions [47].

For model (P_{ref}), the $\mathbf{5PL}$ functions are such that, for a given water flow d_t the $\mathbf{5PL}$ function with the highest value represents the power of the units. As we maximize the profit, for a given water flow, the $\mathbf{5PL}$ with the highest value, for a given d_t, will be considered for the optimal solution if energy prices Λ_t are positive. In this case, the model does not require additional constraints to indicate, for a given t, which one of the variables a_{it} is equal to 1. The considered instances satisfy $\Lambda_t \geq 0$, thus there is no need to add constraints specifying which variable a_{it} is equal to 1 for a time period. Conversely, the additional constraints are required for the PWL model shown in **Section 4.3**. Note that equation (12) is still required, otherwise the model could consider multiple $\mathbf{5PL}$ functions at a given time period, which would induce an incorrect value of the power p_t.

Figure 2b shows the evolution of function $\boldsymbol{g}$ with respect to d_t for instance B-T-1 (see **Table 9** in **Appendix C**). The functions in black are for the minimum and the maximum head, and the grey region represents the power function for the possible values of h_t.



(a) Example of function $\boldsymbol{f}$ (constraint (10))



(b) Example of function $\boldsymbol{g}$ for two units for the minimum and maximum head (constraint (14))

Figure 2: Examples of function $\boldsymbol{f}$ and $\boldsymbol{g}$

Preliminary computations show that (P_{ref}) involves higher computing times than any other model presented, and is not practical for most of the instances considered, even for the smallest ones. This is often the case in real world applications where the functions modeling a physical system are either too complex to be implemented or not supported by any solver.

For solution purposes, the idea is to derive more tractable models than (P_{ref}) to capture the non-linearity in the power function.

4. Models for the 1-HUC

The following sections propose different models to represent the non-linearities of the 1-HUC arising from the power generation function and its characteristics.

The models are also described for the fixed-head 1-HUC, which is a special case often considered.

4.1. (Mixed Integer)Non-Linear Program modeling elementary non-linear functions

The models in this section represent the power function of each unit explicitly, in order to have a representation close to the physics. The downside is that for each unit, auxiliary variables are required. Three models are presented. The first one features a family of polynomial functions, the second one a family of $\mathbf{5}\boldsymbol{PL}$ functions with the **max** function, and the last one a family of $\mathbf{5}\boldsymbol{PL}$ functions without the **max** function.

4.1.1. MINLP with a family of polynomial functions: model ($P_{2D-poly}$)

The power function of one unit for a given head has a parabolic shape. A parabolic shape can be represented with a polynomial function with degree 2. Each polynomial function represents the power generated by a unit, plus the contribution of the previous ones, following their start-up order. **Figure 3** shows an example with two units. We define the following notations:

- b_{it} : the binary variables such that $b_{it} = 1$ if we use the polynomial function of unit i at the time period t;
- y_{kit} : the continuous variables that are the coefficients of monomial d_t^k in the polynomial of unit i at time period t;
- γ_{ki} and δ_{ki} : the constants such that y_{kit} linearly depends on h_t with these parameters.

We introduce the following constraints:

$$y_{kit} = \gamma_{ki} + \delta_{ki} \cdot h_t \qquad \forall k \leq 2, \forall i \leq K, \forall t \leq T \qquad (15)$$

$$\sum_{i=1}^{K} b_{it} = 1 \qquad \forall t \leq T \qquad (16)$$

$$b_{it} \in \{0,1\} \qquad \forall i \leq K, \forall t \leq T \qquad (17)$$

$$p_t = \rho \cdot G \cdot h_t \cdot \sum_{i=1}^{K} b_{it} \cdot \sum_{k=0}^{2} y_{kit} \cdot (d_t)^k \qquad \forall t \leq T \qquad (18)$$

Equalities (15) define parameters y_{kit} as linearly dependent on h_t. Constraints (16) and (17) ensure that only one of the polynomials is active for each time period. Equalities (18) express the power with function $\boldsymbol{g}$ represented by a family of polynomial functions.

The complete model, called ($P_{2D-poly}$) is defined by objective function (2) and constraints (3)-(4), (7)-(8), (10), (15)-(18). It appears that ($P_{2D-poly}$) is a non-convex MINLP as (10) and (18) are non-linear. Indeed, function $\boldsymbol{f}$ computing h_t in (10) is concave, and the polynomial functions in (18) are concave.

But y_{kit} is linear with respect to h_t in (15), and variable y_{kit} is multiplied by h_t in (18). So the power function is convex with respect to h_t, as it is an increasing polynomial of degree 2. Consequently, the region for the optimization is non-convex. This model represents well the power function, as it takes into account characteristics C1, C3 and C4. However, this model still has downsides, mainly the addition of auxiliary binary variables.

In a similar fashion as for model (P_{ref}), model $(P_{2D-poly})$ does not require additional constraints to match the values of b_{it} with the values of d_t. Indeed, for any given water flow d_t, the polynomial with the highest value represents the power of the units. In opposition, one can identify these additional constraints for the PWL model in **Section 4.3**. Note that equation (16) is still required.

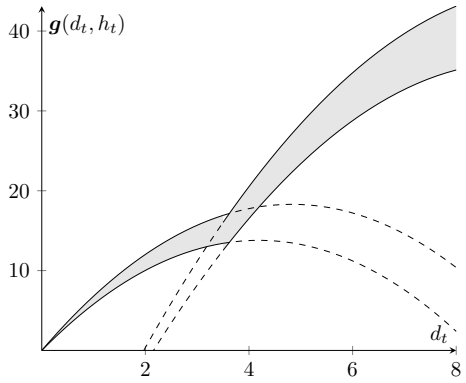


Figure 3: Function $\boldsymbol{g}$ with polynomial functions

Model $(P_{2D-poly})$ for the fixed-head 1-HUC problem. To adapt the model $(P_{2D-poly})$ to the fixed-head 1-HUC problem, with head H, we introduce the following notations:

- Y_{ki} : the constant coefficient of $(d_t)^k$ for unit i.

The power function becomes (19)

$$p_t = \rho \cdot G \cdot h_. \Big(\sum_{i=1}^{K} b_{it} \cdot (\sum_{k=0}^{2} Y_{ki} \cdot (d_t)^k) \Big) \qquad \forall t \leq T \qquad (19)$$

Model $(P_{2D-poly})$ for the fixed-head 1-HUC problem contains objective function (2) and constraints (3)-(4), (7)-(8), (16)-(17), (19) and is an MINLP. Despite the concavity of the polynomials of degree 2 and the fact that we maximize the objective function, the region of optimization for the continuous relaxation is not convex. This is because the polynomials are multiplied by binary variables b_{it}, which become continuous for the continuous relaxation.

4.1.2. NLP with $\mathbf{5PL}$ *functions using function* **max***: model* ($P_{5PL-max}$)

Function $\boldsymbol{g}$ can be represented as a sum of $\mathbf{5PL}$ functions [24], where each $\mathbf{5PL}$ represents the power of one unit. By summing properly parameterized $\mathbf{5PL}$ functions, the sum can be a precise approximation of the physical data. **Figure 4** shows an example of the sum of $\mathbf{5PL}$ functions as a solid line, and the two separated $\mathbf{5PL}$ as dashed lines.

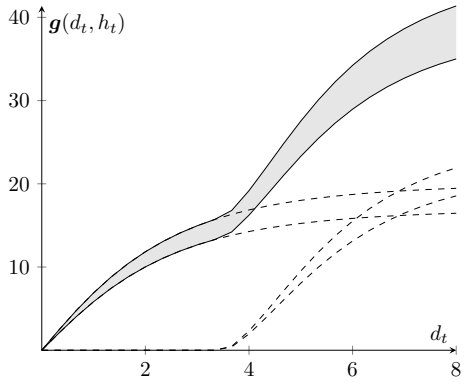


Figure 4: Function $\boldsymbol{g}$ with a sum of $\mathbf{5PL}$ functions

To represent $\boldsymbol{g}$ with such a sum of $\mathbf{5PL}$ functions, the $\mathbf{5PL}$ functions need to depend on the water flow and the head. As such, these parameters are variables and depend on the unit i and time step t. To use $\mathbf{5PL}$ functions, we introduce the following notations:

- z_{jit} : the continuous variables that are the j^{th} parameter of the $\mathbf{5PL}$ function for the unit i at the time period t;
- η_{ji} and θ_{ji} : the constants such that z_{jit} linearly depends on h_t with these parameters.

For this model we need the following equalities:

$$z_{jit} = \eta_{ji} + \theta_{ji} \cdot h_t \qquad \forall j \leq 5, \forall i \leq K, \forall t \leq T \tag{20}$$

$$p_t = \rho \cdot G \cdot h_t \cdot \Big(\sum_{i=1}^{K} z_{4it} + \frac{-z_{4it}}{\Big(1 + \Big(\frac{\mathbf{max}(0, d_t - z_{1it})}{z_{3it}}\Big)^{z_{2it}}\Big)^{z_{5it}}} \Big) \qquad \forall t \leq T \tag{21}$$

Equalities (20) set the parameters of the $\mathbf{5PL}$ functions used in function $\boldsymbol{g}$. Equalities (21) express the power with $\boldsymbol{g}$ as a sum of non-linear functions. These functions are a slight variant of the $\mathbf{5PL}$ functions detailed in **Appendix B**. This is because such $\mathbf{5PL}$ functions are not defined when $d_t < z_{1it}$. Consequently, it is necessary to introduce a **max** function in equalities (21) in order for the $\mathbf{5PL}$ functions to always be defined.

The model ($P_{5PL-max}$) includes objective function (2) and constraints (3)-(4), (7)-(8), (10), (20)-(21). It is a non-convex non-concave NLP, and the non-linearity takes into account characteristics C1 and C2. As this model contains a **max** function, it is not supported by some global optimization solvers.

Model ($P_{5PL-max}$) for the fixed-head 1-HUC problem. To adapt the model ($P_{5PL-max}$) to the fixed-head 1-HUC problem, with head H, we introduce the following notations:

- Z_{ji} : the j^{th} parameters for the $\mathbf{5PL}$ function of the unit i.

The power is defined by equalities (22).

$$p_t = \rho \cdot G \cdot h_. \Big(\sum_{i=1}^{K} Z_{4i} + \frac{-Z_{4i}}{\Big(1 + \Big(\frac{\mathbf{max}(0, d_t - Z_{1i})}{Z_{3i}}\Big)^{Z_{2i}}\Big)^{Z_{5i}}} \Big) \qquad \forall t \leq T \qquad (22)$$

Model ($P_{5PL-max}$) for the fixed-head 1-HUC problem contains objective function (2) and constraints (3)-(4), (7)-(8), (22).

4.1.3. MINLP with $\mathbf{5PL}$ functions using auxiliary variables: model ($P_{5PL-bin}$)

This model is a variation of ($P_{5PL-max}$), where the **max** is linearized by adding linear constraints and auxiliary variables. We introduce the following notations:

- c_{it} : the binary variables equal to 1 if $d_t \geq z_{1it}$;
- m_{it} : the continuous variables equal to $\max(0, d_t - z_{1it})$.

We introduce the following set of constraints:

$$d_t - z_{1it} \leq m_{it} \leq d_t - z_{1it} + (1 - c_{it}) \cdot (\overline{D} - \underline{Z}_{1it}) \qquad \forall i \leq K, \forall t \leq T \qquad (23)$$

$$0 \leq m_{it} \leq c_{it} \cdot (\overline{D} - \underline{Z}_{1it}) \qquad \forall i \leq K, \forall t \leq T \qquad (24)$$

$$c_{it} \in \{0, 1\} \qquad \forall i \leq K, \forall t \leq T \qquad (25)$$

$$p_t = \rho \cdot G \cdot h_t \cdot \Big(\sum_{i=1}^{K} z_{4it} + \frac{-z_{4it}}{\Big(1 + \Big(\frac{m_{it}}{z_{3it}}\Big)^{z_{2it}}\Big)^{z_{5it}}} \Big) \qquad \forall t \leq T \qquad (26)$$

Set of constraints (23)-(25) ensures $m_{it} = \mathbf{max}(0, d_t - z_{1it})$. Equalities (26) express the power in the same manner as equalities (21), but using m_{it}.

The model ($P_{5PL-bin}$) contains objective function (2) and constraints (3)-(4), (7)-(8), (10), (20), (23)-(26). Unlike the NLP model ($P_{5PL-max}$), model ($P_{5PL-bin}$) is an MINLP, as it requires auxiliary binary variable c_{it}. Model ($P_{5PL-bin}$) can be solved by more MINLP solvers than model ($P_{5PL-max}$), as function **max** has been linearized. The representation of the power function is the same for ($P_{5PL-max}$) and ($P_{5PL-bin}$), and both models take into account

characteristics C1 and C2. Note that the model (P_{gen}), with the piecewise non-linear function with $\mathbf{5PL}$, is also an MINLP. The difference is that the binary variables are not the same as the ones in ($P_{5PL-bin}$). Indeed, the binary variables of ($P_{5PL-bin}$) only acts in order to linearize the function **max**, while in (P_{gen}) they are decision variables.

Model ($P_{5PL-bin}$) for the fixed-head 1-HUC problem. To adapt the model ($P_{5PL-bin}$) to the fixed-head 1-HUC problem, with head H, we introduce the following notations:

- Z_{ji} : the constants for the parameter j for the $\mathbf{5PL}$ function of the unit i;
- m_{it} : the continuous variables such that $m_{it} = \mathbf{max}(0, d_t - Z_{1i})$.

To ensure the behaviour of variable m_{it}, we add the set of constraints (27)-(30), defined as (23)-(26), where variables z_{jit} are replaced by constants Z_{ji} for all $t \leq T$. Model ($P_{5PL-bin}$) for the fixed-head 1-HUC problem contains objective function (2) and constraints (3)-(4), (7)-(8), (27)-(30).

4.2. (Mixed Integer)Non-Linear Program modeling an aggregated non-linear function

The models introduced in this section represent all units as an aggregated function. The principle is to consider a single function to represent the whole power function, instead of a family or a sum of elementary functions. A single function being less precise, the expected benefit is a quick solution by MINLP tools, as few additional variables and constraints are required. The functions we propose are the following: a polynomial function, a bilinear function, and a finite set of operating flows.

4.2.1. NLP with a high degree polynomial function: model ($P_{HD-poly}$)

A model using an aggregated function that represents well the physics is obtained by using a single polynomial function as function $\boldsymbol{g}$. **Figure 5** shows an example of an 8^{th} degree polynomial function for an instance with two units. As $\boldsymbol{g}$ depends on the head h_t, the coefficients of the polynomial are linearly dependent on h_t. We introduce the following notation

- Q : the degree of the polynomial, with $Q = 4 \cdot K$, where K denotes the number of units;
- u_{qt} : the continuous variable that are the coefficient of monomial d_t^q in the polynomial function at time period t;
- μ_q and ν_q : the constants such that u_{qt} linearly depends on h_t with these parameters.

For this model, we need the following equalities:

$$u_{qt} = \mu_q + \nu_q \cdot h_t \qquad \forall q \leq Q, \forall t \leq T \qquad (31)$$

$$p_t = \rho \cdot G \cdot h_t \cdot \Big(\sum_{q=0}^{Q} u_{qt} \cdot (d_t)^q \Big) \qquad \forall t \leq T \qquad (32)$$

Equalities (31) set the parameters of the polynomial function. Equalities (32) define the power with $\boldsymbol{g}$ as a polynomial function.

The model, called $(P_{HD-poly})$ includes objective function (2) and constraints (3)-(4), (7)-(8), (10), (31)-(32). It is an NLP featuring characteristic C1 as it is non-convex and non-concave. The benefits compared to $(P_{2D-poly})$ is that $(P_{HD-poly})$ considers only a single polynomial function. This means that no auxiliary binary variables are required. The downside of $(P_{HD-poly})$ is that high degree polynomials (8 for two units, 20 for five units) can induce large approximation errors. If the water flow can fluctuate between 0 and 100, then it means that the solver might need to compute numbers such as 0.1^8 or 100^8, which are either too small or too large numbers for standard solvers' precision. Moreover, computational errors can have a dramatic impact for the HUC, as an error for a time period will cumulate and carry over all future time periods [44].

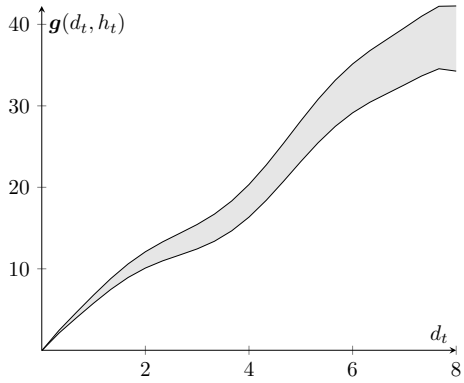


Figure 5: Function $\boldsymbol{g}$ with a single polynomial function

Model $(P_{HD-poly})$ for the fixed-head 1-HUC problem. To adapt the model $(P_{HD-poly})$ to the fixed-head 1-HUC problem, with head H, we introduce the following notations:

- U_q : the coefficients for the degree q of the polynomial function.

The power function becomes:

$$p_t = \rho \cdot G \cdot h. \Big(\sum_{q=0}^{Q} U_q \cdot (d_t)^q \Big) \qquad \forall t \leq T \qquad (33)$$

The model $(P_{HD-poly})$ for these special instances contains objective function (2) and constraints (3)-(4), (7)-(8), (33).

4.2.2. NLP with a bilinear function: model (P_{bilin})

A type of model often described in the literature to solve the HUC as an MINLP is a bilinear model [32], [8]. **Figure 6** shows an example of a bilinear function. The power is linear with respect to the water flow, and to the head. In the model (P_{gen}), the power is already linear with respect to the head. We need to make it also linear with respect to the water flow. To do so, we introduce the following notations:

- ϕ and ψ: the constants such that the power is linearly dependent on the water flow.

We adapt the power function as follows:

$$p_t = \rho \cdot G \cdot h_t \cdot (\phi + \psi \cdot d_t) \qquad \forall t \leq T \tag{34}$$

Equalities (34) express the power as a bilinear function of the head h_t and the water flow d_t.

The model (P_{bilin}) contains objective function (2) and constraints (3)-(4), (7)-(8), (10), (34). Equalities (34) feature a bilinear function of h_t and d_t. A bilinear function is non-convex non-concave even if both variables are positive [40]. However, this function remains simpler than the ones featured in previously described models, such as polynomial or $\mathbf{5PL}$ functions. Besides, unlike ($P_{2D-poly}$), ($P_{HD-poly}$), ($P_{5PL-max}$) or ($P_{5PL-bin}$), model (P_{bilin}) does not require any additional binary variables. This makes this model a potential candidate to solve quickly the problem. The downside is that this model has the roughest approximation of all models. Indeed, the bilinear function features none of the non-linear characteristics C1, C2, C3 or C4.

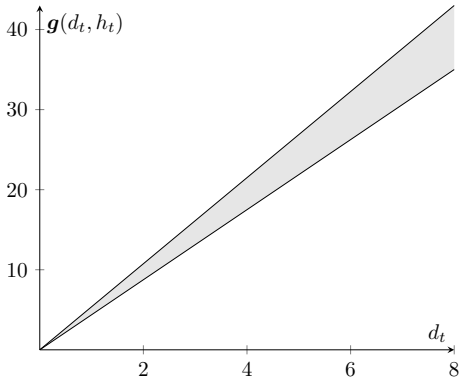


Figure 6: Function $\boldsymbol{g}$ with a bilinear function

Model (P_{bilin}) for the fixed-head 1-HUC problem. When we adapt the model (P_{bilin}) to the fixed-head 1-HUC problem, with head H, the model becomes a linear model, where the power is a linear function of the water flow. To do so,

we simply adapt the power function as follows:

$$p_t = \rho \cdot G \cdot h_{\cdot}(\phi + \psi \cdot d_t) \qquad \forall t \leq T \tag{35}$$

The model (P_{bilin}) for the fixed-head 1-HUC problem contains objective function (2) and constraints (3)-(4), (7)-(8), (35), which yields an LP.

4.2.3. MINLP with a discrete set of decisions: model (P_{op})

Having a discrete set of decisions for the 1-HUC problem means that only a given number of operating flows, say L, are authorized. **Figure 7** shows an example of a discrete set of decisions, with $L = 9$. These operating flows are specifically chosen where the power production is the most profitable, and usually are in the concave parts of the original power function in (P_{ref}). We introduce the following notations:

- d_i : the constant being the i^{th} operating flow;
- o_{it} : the binary variable such that $o_{it} = 1$ if we use the i^{th} operating flow at time period t.

We consider a model with disjunctive constraints between the operating flows. As such, we need the following constraints:

$$\sum_{i=1}^{L} o_{it} \leq 1 \qquad \forall t \leq T \tag{36}$$

$$v^1_{t'} = V^1_0 + \sum_{t=1}^{t'} \Big(A^1_t - (\sum_{i=1}^{L} o_{it} \cdot d_i) \cdot \Delta \Big) \qquad \forall t' \leq T \tag{37}$$

$$v^2_{t'} = V^2_0 + \sum_{t=1}^{t'} \Big(A^2_t + (\sum_{i=1}^{L} o_{it} \cdot d_i) \cdot \Delta \Big) \qquad \forall t' \leq T \tag{38}$$

$$p_t = \rho \cdot G \cdot h_t \cdot \boldsymbol{g}(\sum_{i=1}^{L} o_{it} \cdot d_i, h_t) \qquad \forall t \leq T \tag{39}$$

Inequalities (36) ensure that only one operating flow can be active at each time period. The set of equalities (37)-(39) corresponds to equalities (3), (4) and (6) from (P_{gen}), with operating flows instead of the water flow d_t.

This leads to a new generic model (P_{op}) containing objective function (2) and constraints (7), (10), (36)-(39). Function $\boldsymbol{g}$ in (39) can be any of the previously described function for models $(P_{HD-poly})$, $(P_{2D-poly})$, $(P_{5PL-max})$, $(P_{5PL-bin})$ or (P_{bilin}). Because we have a finite set of operating flows, function $\boldsymbol{g}$ for (P_{op}) will feature none of the characteristics C1 to C4, regardless of the function considered. Model (P_{op}) can be beneficial because its solution space is drastically smaller than the others but does not offer as much freedom, in particular when target volumes occur. As the operating flows are amongst the most profitable ones, the solution might still be close to the optimal solution. The downside of this model is that target volumes can be unreachable with the chosen set of operating flows, thus leading to infeasibility.

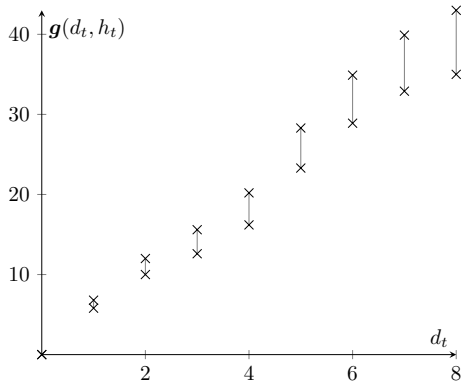


Figure 7: Function $\boldsymbol{g}$ with a discrete set of decisions

Model (P_{op}) for the fixed-head 1-HUC problem. When we adapt the model (P_{op}) for the fixed-head 1-HUC problem, with head H, the model becomes a MILP model. In this case the power only depends on the water flow, thus there is a finite set of possible powers. As such, the model becomes a MILP as we have to choose a pair (operating flow, power produced) amongst a list of pairs at each time period. We introduce the following notations:

- p_i : the constant being the power generated for the i^{th} operating flow.

The model is very similar to (P_{op}) for the general 1-HUC problem, but we define the power differently as we use constants p_l:

$$p_t = \sum_{i=1}^{L} o_{it} \cdot p_i \qquad \forall t \leq T \tag{40}$$

The model (P_{op}) adapted for the fixed-head 1-HUC problem contains objective function (2) and constraints (7), (36)-(38), (40). We can notice that this model contains very few variables, as only the decision variables o_{it} are required.

4.3. Mixed Integer Linear Program modeling piecewise linear functions: model (P_{pwl})

The model of this section follows the common practice of using a PWL approximation when modeling a non-linear expression. **Figure 8** shows an example of a piecewise linear function. Our comparisons of a PWL model with the previously described non-linear models are mostly focused on the precision, i.e., the quality of the solutions obtained. As such, we will consider a standard PWL formulation [11] [22], more precisely the convex combination formulation. There exist much more efficient formulations, e.g. the logarithmic formulation in [51], but they will not be considered as it will not impact the value of the solution, but only the computing time.

A generic way to obtain a two-dimensional PWL function is to use the one-dimensional method described in [14]. It is a generalization of the convex combination formulation [22]. The method described to approximate a non-linear function $\boldsymbol{f}(x, y)$ is as follows. We fix I variables on the x axis, $(\widetilde{x}_1, .., \widetilde{x}_I)$, and J variables on the y axis $(\widetilde{y}_1, .., \widetilde{y}_J)$. For each $\widetilde{y}_j$, we approximate $f(x, \widetilde{y}_j)$ with a PWL function $\boldsymbol{l}(x, \widetilde{y}_j)$, where each $\widetilde{x}_i$, $i \leq I$ acts like a break point. It means that piece i of $\boldsymbol{l}(x, y_j)$ is a linear function between $\widetilde{x}_i$ and $\widetilde{x}_{i+1}$. We obtain then J PWL functions with $I - 1$ pieces. The value for $\boldsymbol{l}(x, y)$, $y \in [\widetilde{y}_j, \widetilde{y}_{j+1}]$, is approximated by $\boldsymbol{l}(x, \widetilde{y}_j)$. For the 1-HUC problem, we will approximate the power function with respect to the water flow d_t for a set of fixed volumes $\widetilde{y}_j$, $j \in 1, ..., J$.

In this model, we aggregate both non-linear functions $\boldsymbol{f}$ and $\boldsymbol{g}$ as a unique function to represent the power. To do so, it is possible to replace h_t by $f(v_t^1)$ in equalities (6) from (P_{gen}). Thus, the power is defined as follows, and we only need to approximate one two-dimensional non-linear function for the whole model:

$$p_t = \boldsymbol{F}(d_t, v_t^1) = \rho \cdot G \cdot \boldsymbol{f}(v_t^1) \cdot \boldsymbol{g}(d_t, \boldsymbol{f}(v_t^1))$$

To use the PWL approximation, we introduce the following notations:

- $\boldsymbol{l}(v_t^1, d_t)$: the PWL approximation of $\boldsymbol{F}(v_t^1, d_t)$;
- J : the number of PWL functions allocated on the volume axis;
- I : the number of breakpoints on the water flow axis;
- $\tilde{v}_j^1$: the volume corresponding to the j^{th} one dimensional PWL function;
- $\tilde{d}_i$: the breakpoint i for all J PWL functions.

And we include the following variables:

- $l_{j,t}$: the value $\boldsymbol{l}(\tilde{v}_j^1, d_t)$ of the PWL function j at time period t;
- r_{it} : the binary variables such that $r_{it} = 1$ if d_t is located on the interval $[\tilde{d}_i, \tilde{d}_{i+1}]$;
- w_{it} : the continuous variables such that d_t is the convex combination $w_{it}\tilde{d}_i + w_{i+1t}\tilde{d}_{i+1}$;
- s_{jt} : the binary variables such that $s_{jt} = 1$ if v_t^1 is located in the interval $[\tilde{v}_j^1, \tilde{v}_{j+1}^1]$.

The PWL formulation requires the following constraints:

$$\sum_{i=1}^{I} r_{it} = 1 \qquad \forall t \leq T \qquad (41)$$

$$w_{it} \leq r_{i-1t} + r_{it} \qquad \forall i \leq I, \forall t \leq T \qquad (42)$$

$$\sum_{i=1}^{I} w_{it} = 1 \qquad \forall t \leq T \qquad (43)$$

$$d_t = \sum_{i=1}^{I} w_{it} \cdot \tilde{d}_i \qquad \forall t \leq T \qquad (44)$$

$$l_{j,t} = \sum_{i=1}^{I} w_{it} \cdot \boldsymbol{F}(\tilde{v}_j^1, \tilde{d}_i) \qquad \forall j \leq J, \forall t \leq T \qquad (45)$$

$$r_{it} \in \{0, 1\} \qquad \forall i \leq I, \forall t \leq T \qquad (46)$$

$$0 \leq w_{it} \leq 1 \qquad \forall i \leq I, \forall t \leq T \qquad (47)$$

$$\sum_{j=1}^{J} s_{jt} \cdot \tilde{v}_j^1 \leq v_t^1 \leq \sum_{j=1}^{J} s_{jt} \cdot \tilde{v}_{j+1}^1 \qquad \forall t \leq T \qquad (48)$$

$$\sum_{j=1}^{J} s_{jt} = 1 \qquad \forall t \leq T \qquad (49)$$

$$l_{j,t} - \overline{P}_t \cdot (1 - s_{jt}) \leq p_t \leq l_{j,t} + \overline{P}_t \cdot (1 - s_{jt}) \qquad \forall j \leq J, \forall t \leq T \qquad (50)$$

$$s_{jt} \in \{0, 1\} \qquad \forall j \leq J, \forall t \leq T \qquad (51)$$

Constraints (41)-(47) are the standard convex combination formulation for a one-dimensional PWL function, applied to approximate function $\boldsymbol{F}$ for each given volume. These constraints ensure that $l_{j,t}$ is the value, at time period t of the PWL function approximating $\boldsymbol{F}$ for volume $\tilde{v}_j^1$. Equalities (41) express that exactly one variable r_{it} is equal to one at time period t, meaning that we consider one piece of the PWL function at time period t. Inequalities (42) allow the weight r_{it} of a breakpoint to be greater than zero only if one for the two surrounding pieces is considered at time period t. Equalities (43)-(44) ensure that d_t is the convex combination of the $\tilde{d}_i$ with the weights r_{it} at time period t. Equalities (45) define the PWL approximation of the power function. Constraints (46)-(47) give the domain of variables r_{it} and w_{it}.

We have described the constraints to obtain $l_{t,j}$ the approximated value of $\boldsymbol{F}$ at time period t using univariate PWL functions. Now we need constraints (48)-(51) in order to obtain the power from the value $l_{t,j}$ for the fixed volume $\tilde{v}_j^1$, with $\tilde{v}_j^1 \leq v_t^1 \leq \tilde{v}_{j+1}^1$. Constraints (48)-(49) ensure $s_{j,t} = 1$ if $v_t^1 \in [\tilde{v}_j^1; \tilde{v}_{j+1}^1]$, and exactly one variable $s_{j,t}$ is equal to 1. Variable $s_{j,t}$ then indicates which PWL function should be considered depending on the volume. Inequalities (50) ensure $p_t = l_{j,t}$ if $s_{j,t} = 1$, or give trivial bounds if $s_{j,t} = 0$. Constraints (51) provide the domain of s_{jt}.

The MILP model (P_{pwl}) contains objective function (2) and constraints (3)-(4), (7)-(8), (41)-(51). The consequences of this model being a MILP are twofold. On the one hand, it can be solved with powerful MILP tools. On the other hand it includes a lot of auxiliary variables and constraints, and it does not include any of the non-linear characteristics of the power function.

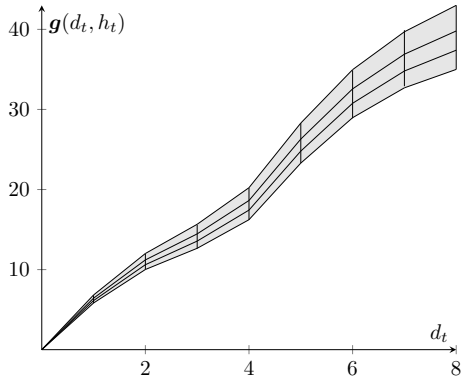


Figure 8: Function $\boldsymbol{g}$ with a piecewise-linear function

Model (P_{pwl}) for the fixed-head 1-HUC problem. When we adapt the model (P_{pwl}) to the fixed-head 1-HUC problem, with head H, we approximate a one-dimensional power function. To do so, we use the convex combination formulation [22], which is the formulation generalized for model (P_{pwl}) in the general case. The convex combination formulation adapted for the fixed-head 1-HUC problem requires constraints (41)-(44) and (46)-(47). We express the power as follows:

$$p_t = \sum_{i=1}^{I} w_{it} \cdot \rho \cdot G \cdot h.\boldsymbol{g}(\tilde{d}_i, H) \qquad \forall t \leq T \qquad (52)$$

Thus the model (P_{pwl}) for these special instances contains objective function (2) and constraints (3)-(4), (7)-(8), (41)-(44),(46)-(47), (52).

4.4. Summary of models and non-linear functions

We have described a total of fourteen different models. All models have the same objective function (2). Also, most of the models share the same set of constraints. We define constraint sets S1=(3)-(4),(7)-(8),(10) and S2=(3)-(4),(7)-(8). **Table 2** summarizes all models with their constraints. The difference between these models is the representation of the power function.

Table 3 shows the characteristics and the type of program for each model. The convexity and the linearity of a model do not take into account the integer variables, thus concerns the continuous relaxation. From a theoretical point of view, none of the presented models perfectly fits the power function of (P_{ref}).

Table 2: Summary of the proposed models

Model	Section	General cases	Fixed-head cases
$(P_{2D-poly})$	4.1.1	S1,(15)-(18)	S2,(16)-(17),(19)
$(P_{5PL-max})$	4.1.2	S1,(20)-(21)	S2,(22)
$(P_{5PL-bin})$	4.1.3	S1,(20),(23)-(26)	S2,(27)-(30)
$(P_{HD-poly})$	4.2.1	S1,(31)-(32)	S2,(33)
(P_{bilin})	4.2.2	S1,(34)	S2,(35)
(P_{op})	4.2.3	(2),(7),(10),(36)-(39)	(2),(7), (36)-(38),(40)
(P_{pwl})	4.3	S2,(41)-(51)	S2,(41)-(44),(46)-47,52

Indeed, none of the models features all four non-linear characteristics of the power functions. However, it will be shown in the numerical experiments that some models allow for very small approximation errors, while other models, with simpler non-linear expressions, lead to shorter computing times. **Table 4** shows the size of each model, namely the number of constraints (#cst), the number of binary and continuous variables (respectively #b-var and #c-bar). We recall that N is the number of units, T the number of time periods, Q the degree of the polynomial function in model $(P_{HD-poly})$, I the number of breakpoints of the PWL functions and J the number of PWL functions in model (P_{pwl}).

Table 3: Comparison of the models non-linear characteristics

	1-HUC		Fixed-head 1-HUC		Characteristics			
Model	Type	Convexity	Type	Convexity	C1	C2	C3	C4
$(P_{2D-poly})$	MINLP	non-convex	MINLP	non-convex	✓	✗	✓	✓
$(P_{5PL-max})$	NLP	non-convex	NLP	non-convex	✓	✓	✗	✗
$(P_{5PL-bin})$	MINLP	non-convex	MINLP	non-convex	✓	✓	✗	✗
$(P_{HD-poly})$	NLP	non-convex	NLP	non-convex	✓	✗	✓	✗
(P_{bilin})	NLP	non-convex	LP	linear	✗	✗	✗	✗
(P_{op})	MINLP	non-convex	MILP	linear	✗	✗	✗	✗
(P_{pwl})	MILP	linear	MILP	linear	✗	✗	✗	✗

Table 4: Nomber of constraints, binary variables and continuous variables for each model

	1-HUC			Fixed-head 1-HUC		
Model	#cst	#b-var	#c-var	#cst	#b-var	#c-var
$(P_{2D-poly})$	$(11+3\cdot N)\cdot T$	$N\cdot T$	$(5+3\cdot N)\cdot T$	$10\cdot T$	$N\cdot T$	$4\cdot T$
$(P_{5PL-max})$	$(10+5\cdot N)\cdot T$	0	$(5+5\cdot N)\cdot T$	$9\cdot T$	0	$4\cdot T$
$(P_{5PL-bin})$	$(10+15\cdot N)\cdot T$	$N\cdot T$	$(5+10\cdot N)\cdot T$	$(9+10\cdot N)\cdot T$	$N\cdot T$	$(4+5\cdot N)\cdot T$
$(P_{HD-poly})$	$(10+Q)\cdot T$	0	$(5+Q)\cdot T$	$9\cdot T$	0	$4\cdot T$
(P_{bilin})	$10\cdot T$	0	$5\cdot T$	$9\cdot T$	0	$4\cdot T$
(P_{op})	$9\cdot T$	$L\cdot T$	$5\cdot T$	$8\cdot T$	$L\cdot T$	$4\cdot T$
(P_{pwl})	$(14+4\cdot I+4\cdot J)\cdot T$	$(I+J)\cdot T$	$(4+I+J)\cdot T$	$(12+4\cdot I)\cdot T$	$I\cdot T$	$(4+I)\cdot T$

It is also possible to compare the models, on the basis of the difficulty for the solvers to handle their non-linear expression. A way to measure this is to compare the size of the reformulation binary tree for the non-linear expressions

as in [48]. Following this metric, $\mathbf{5PL}$ functions are by far the most difficult functions, followed by high degree polynomials, two degree polynomials, bilinear functions and linear functions.

Recall that the power function of model ($P_{5PL-max}$) features a max function. Consequently, the power function for ($P_{5PL-max}$) is nonsmooth and with discontinuous derivatives. This means that this model is not supported by all non-linear solvers.

5. Numerical experiments

The computational evaluation is performed via Neos Server [12] in October and November 2021. All experiments are performed on Neos Server machine prod-exec-7 (a 2x Intel Xeon Gold 5218 @ 2.3GHz processor with 384 GB of RAM), using a single thread. The computing time limit is set to 10800 seconds.

5.1. Solvers considered

Five non-linear solvers are considered in this numerical comparison, ANTIGONE [33], BARON [49], COUENNE [5], LINDOGlobal [31] and SCIP [52], as they are accessible on Neos Server [12] with GAMS format. All of the five solvers implement global optimization algorithms, and use the sBB algorithm or its derivatives. To improve the solution time and quality, the solvers complement variations of the sBB with different tools. However, the set of implemented tools greatly differs from one solver to another. As shown in the following, it can happen that one solver is more efficient than another one for a given model, while the opposite can be true for another model. This variation is the reason why we consider multiple non-linear solvers, in order not to have a biased conclusion due to the solvers.

As comparing linear solvers is beyond the scope of this work, a well known and efficient linear solver CPLEX [10] is arbitrarily chosen to solve MILP models. Solver CPLEX is accessible on Neos Server with LP format.

The six solvers are further described in **Appendix A**.

5.2. Instances, parameters, terminology and metrics

Parameters of the original power function. The numerical values for the parameters of the power functions featured in the different models are obtained by fitting their power functions to the one of (P_{ref}). The fitting is done via Scipy's curve_fit function[1], using a non-linear least squares method. This approach does not provide an a priori precision for the resulting function with respect to the data. Recall that the purpose of this work is to study and analyse various approximations of the power function. Thus the parameters of the head functions of all proposed models are the ones of (P_{ref}).

[1] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html`, accessed: 2023-01-09

Parameters of model (P_{pwl}). For model (P_{pwl}) we also want to take into account the impact of the number of linear pieces, therefore we will define three variants of (P_{pwl}): (P^1_{pwl}), (P^2_{pwl}) and (P^3_{pwl}), respectively with I and J equal to 5, 20 and 100. For every variant, the breakpoints are defined as equidistant instead of being tailored to each instance. Preliminary results show that model (P_{pwl}) is not significantly penalized in terms of approximation error compared to the other models, with equidistant breakpoints. Hence, we will not investigate the best distribution of the breakpoints, in order to keep a simple model for each modeling alternative.

Parameters of model (P_{op}). For model (P_{op}), a discrete set of decision variables is to be chosen. For the considered instances, we define $\boldsymbol{g}$ as the non-linear function of model ($P_{5PL-bin}$), and we consider 5 operating flows per unit. We will not consider models with more operating flows, as the model contains $\mathbf{5PL}$ functions that are already difficult to handle. Additional operating flows would make the model irrelevant as it would become too hard to solve.

Variable bounds. All models contain variables that are subject to an equality constraint such as the head h_t or power p_t. These variables have physical bounds, but these are not expressed explicitly in the models, as they are set through the equality constraints. However, when using global solvers, it is a good practice to bound every variable. Hence, for the experiments, every variable has an upper and lower bound, even if these bounds are trivially satisfied through the equality constraints. Note also that these bounds could be improved, but this requires a complete work which is out of the scope of this study. Moreover, solvers usually implement bound tightening techniques. As we also compare the solvers for the proposed model, we do not implement extra features (reformulation, bound tightening etc...). As such the solvers are compared taking into account their complete sets of tools.

Instances. All instances considered are derived from parameter sets A and B, detailed in **Appendix C**. These parameter sets are inspired from real instances from EDF. Different variants of these sets are created to form a larger set of instances. The varying non-linear features and the corresponding parameters of the 1-HUC problem are as follows. The *size of the instance* varies with the number of time periods. *Equality constraints* appear as soon as target volumes are accounted for in the instance. Two features of the non-linear function can be changed: the *number of inflection points* and the *degree of non-linearity*. These features are respectively linked to the number of units, and to when the transition from a unit to another occurs when increasing or decreasing the water flow. The last feature is the *sensitivity of the decision variables*, which measures how much the decision can affect the dynamical behavior of the physical system. For the 1-HUC problem, the smaller the water flows are relative to the absolute volume, the less the volumes change over the time periods. In **Appendix C**, the features are further described, and an equation for the sensitivity of the decision variables is provided. Also, **Table 13** summarizes the instances and their features.

Terminology, notations and metrics. We introduce additional terminology to compare the different models on the considered instances. For the comparisons, we also define the metrics used and their notation.

A *configuration* is defined as a pair (instance, model). For all models, a configuration is solved to optimality when the optimality gap between the primal and dual values is below 0.1%. Note that the optimality gap is not computed in the exact same manner for every solver, but remains very similar. The *optimal solution of a configuration* is the solution when the configuration is solved to optimality. The value of a configuration is the value of its optimal solution. The *recalculated value of a solution* is the value of the solution, evaluated with the non-linear functions of (P_{ref}). Such value is obtained as follows. Consider the water flows of a solution. The recalculated value of this solution is the value of the objective function of (P_{ref}) with the same water flows. The *recalculated value of a configuration* is the recalculated value of the optimal solution of the configuration. A configuration is *solvable* by a solver if the solver supports the model. A configuration is *solved* if it is solved to optimality with at least one solver within the time limit. A configuration is *feasible* with a solver when it is not solved to optimality, but a solution is found within the time limit. A configuration is *infeasible* with a solver if the solver proves the configuration to have no feasible solution within the time limit.

The metrics used to compare the models and the solvers are as follows. The *computing time* (CT) of a configuration is the time required to return the optimal solution. The *approximation error* (AE) of a configuration is the relative difference between the value of the optimal solution of the configuration, and the recalculated value of the configuration. The *distance to the best recalculated value* (DB) of a configuration is the relative difference between the recalculated value of the configuration, and the highest recalculated value of all configurations with the same instance.

As specified, configurations are solved with several solvers. We define the *virtual best solver* (VBS) [27] of a given configuration as the solver that requires minimal CT to solve the configuration. Results show that the AE (resp. the DB) of a configuration is the same for every solver. Thus, for our results, the VBS is the solver that has the configuration solved to optimality in minimal CT. For our analysis we use the metrics of the configurations with their VBS. All figures and tables for the results are with the VBS, except for **Tables 8** and **9** that display the results for each solver. Note that some configurations are not solvable with every solver. Indeed, model ($P_{5PL-max}$) is only supported by LINDOGlobal and SCIP.

5.3. Model comparison

In this section, we present the results of the model comparison.

Results summary. To summarize the results, **Figure 9** shows a bargraph which represents two categories of results. First, the height of the bar for a model corresponds to the proportion of configurations solved with their VBS. Second, the lightest color shows for a model the proportion of configurations for which

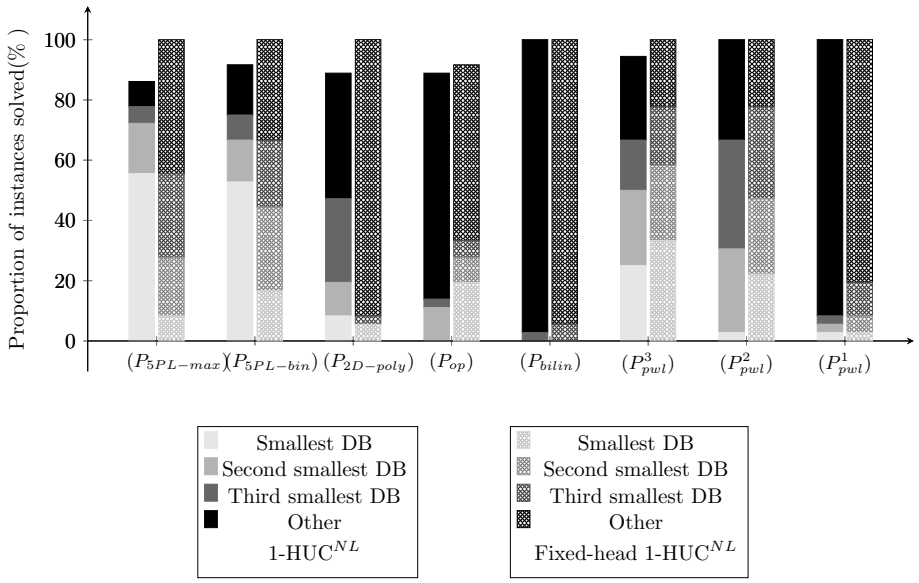


Figure 9: Proportion of configurations for each model solved by their VBS

the model has the lowest DB compared to the other models. Similarly, the second and third lightest color for a model correspond to the proportion of configurations for which the model has the second and third lowest DB. The darkest color for a model corresponds to the proportion of configurations for which the model does not yield a DB which is amongst the three lowest DB. These results are distinguished for both the 1-HUC problem and the fixed-head 1-HUC problem.

In the remainder of this section, we highlight key observations of the model comparison for both 1-HUC problems, with and without a fixed-head.

Finding 1: Infeasibility of model $(P_{HD-poly})$. None of the configurations with $(P_{HD-poly})$ returns a feasible solution.

Model $(P_{HD-poly})$ contains high degree polynomials which can yield very large and small numbers. This produces floating point errors for the solvers, which makes this model impractical without a dedicated solver.

It follows that results related to model $(P_{HD-poly})$ are not enclosed.

Finding 2: Models (P_{op}), $(P_{2D-poly})$ *can yield infeasibilities.* **Table 5** shows, for each model, the proportion of solved (%solved), feasible (%feasible), and infeasible (%infeasible) configurations with their VBS. Note that there is no case where the status is undefined: for every configuration, either a feasible solution is found, or the infeasibility is proven within three hours.

In the case of (P_{op}) infeasibilities can happen because there are instances with target volumes that cannot be reached with the finite set of operating

flows. In the case of $(P_{2D-poly})$, infeasibilities occurs for instances of the 1-HUC problem with high degree of non-linearity. However, it is not clear why such configurations are deemed infeasible by the solvers.

Table 5: Proportion of solution status returned for the configurations for each model by their VBS

	1-HUC			Fixed-head 1-HUC		
Model	%solved	%feasible	%infeasible	%solved	%feasible	%infeasible
$(P_{5PL-max})$	86.1	13.9	0.0	100.0	0.0	0.0
$(P_{5PL-bin})$	91.7	8.3	0.0	100.0	0.0	0.0
$(P_{2D-poly})$	88.9	0.0	11.1	100.0	0.0	0.0
(P_{op})	88.9	2.8	8.3	91.7	0.0	8.3
(P_{bilin})	100.0	0.0	0.0	100.0	0.0	0.0
(P^3_{pwl})	94.4	5.6	0.0	100.0	0.0	0.0
(P^2_{pwl})	100.0	0.0	0.0	100.0	0.0	0.0
(P^1_{pwl})	100.0	0.0	0.0	100.0	0.0	0.0

Finding 3: Considering a fixed-head leads to reduced CT. **Figure 10** shows on the y-axis the proportion of configurations solved with their VBS, under a CT threshold given on the x-axis. The configurations are color-coded depending on the model of the configuration.

Clearly, for every model, the CT is reduced when solving the fixed-head 1-HUC problem.

There are multiple reasons why the CT is lower for the fixed-head case. First, a single one-dimensional non-linearity is considered, in opposition to a two-dimensional one and a one-dimensional one in the general case. Also, for all models by (P_{bilin}), fewer variables are required when the head is fixed. Then, models (P_{op}), (P_{bilin}) have a linear relaxation in the fixed-head case, whereas these models have a non-linear relaxation in the general case.

Finding 4: Considering a fixed-head leads to increased AE, yielding similar AE for all models. **Figure 11** shows on the y-axis the proportion of configurations solved with their VBS, under an AE threshold given on the x-axis. The configurations are color-coded depending on the model of the configuration.

It appears that the AE increases for each model in the fixed-head case. In addition, all models then yield a very similar AE.

We can explain the increased AE for the fixed-head case as follows. In practice, when the volume changes, then the head also changes. However this is not captured when the fixed-head is considered. As such, a fixed-head model will induce higher AE independently of the model selected. This is why all models yield high and similar AE.

Finding 5: Models (P_{bilin}), (P^2_{pwl}), $(P_{2D-poly})$ and $(P_{5PL-bin})$ are the most efficient ones for the 1-HUC problem. **Table 6** shows for each model, the average CT and AE for both the 1-HUC problem and the fixed-head 1-HUC problem. When calculating the average CT, a CT of 10800 seconds (the time limit) is

considered for configurations that are not solved to optimality. Infeasible configurations are not taken into account for this average. **Figure 12** depicts the trade-off between the average CT and AE for each model as a Pareto-front, for the 1-HUC problem with and without fixed-head.

When considering the CT and the AE metrics as two criterias, models (P_{bilin}), (P^2_{pwl}), $(P_{2D-poly})$ and $(P_{5PL-bin})$ are the most efficient ones. Indeed, for any other model, one of the four cited models has a lower CT and a lower AE. We cannot deduce the best overall model, as it depends on the user needs.

Finding 6: Models (P^2_{pwl}), $(P_{2D-poly})$, (P_{op}) and $(P_{5PL-max})$ are the most efficient ones for the fixed-head 1-HUC problem. In a similar fashion as for the previous finding, we detect models (P^2_{pwl}), $(P_{2D-poly})$, (P_{op}) and $(P_{5PL-max})$ to be the most efficient ones. Note that models (P^2_{pwl}), $(P_{2D-poly})$ are amongst the most efficient for the 1-HUC problem with and without a fixed-head.

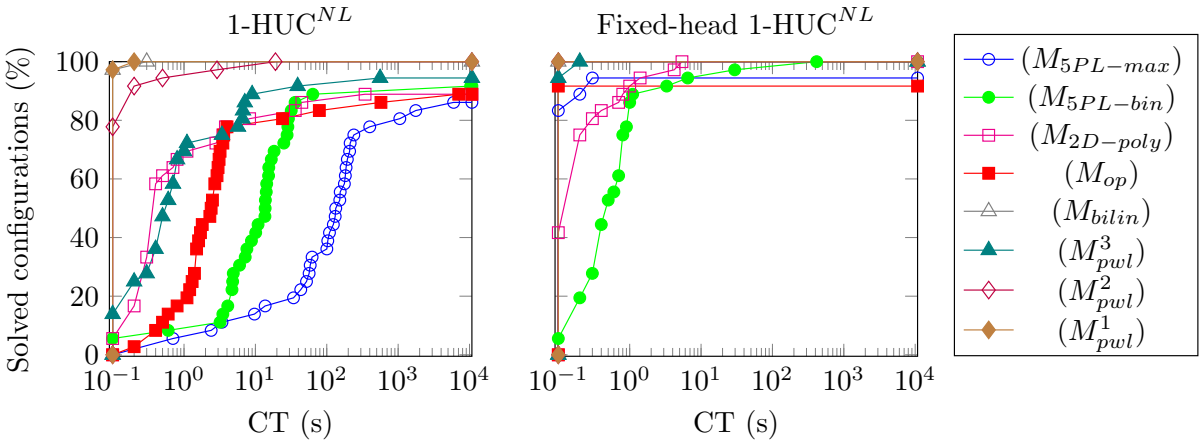


Figure 10: Proportion of configurations solved with their VBS where the CT is under a CT threshold

Table 6: Average CT and AE for each model with its VBS, and CT and AE trade-off for each pair of models for the 1-HUC problem

	model	$(P_{5PL-max})$	$(P_{5PL-bin})$	$(P_{2D-poly})$	(P_{op})	(P_{bilin})	(P^3_{pwl})	(P^2_{pwl})	(P^1_{pwl})
1-HUC	average CT	1832.1	1200.6	12.6	520.1	0.1	618.4	0.7	0.1
	average AE	1.16	0.96	2.88	4.16	19.14	3.43	7.33	40.26
fixed-head 1-HUC	average CT	600.6	13.1	0.5	3.3	0.1	0.1	0.1	0.1
	average AE	12.90	13.85	14.11	13.43	37.92	18.34	18.14	18.93

Finding 7: Metrics DB and AE are correlated. **Figure 13** shows on the y-axis the proportion of configurations solved with their VBS, under a DB threshold given on the x-axis.

Results show that $(P_{5PL-bin})$ and $(P_{5PL-max})$ solve more than 60% of the 1-HUC problem instances with a DB below 0.01%. In opposition, models (P^1_{pwl})

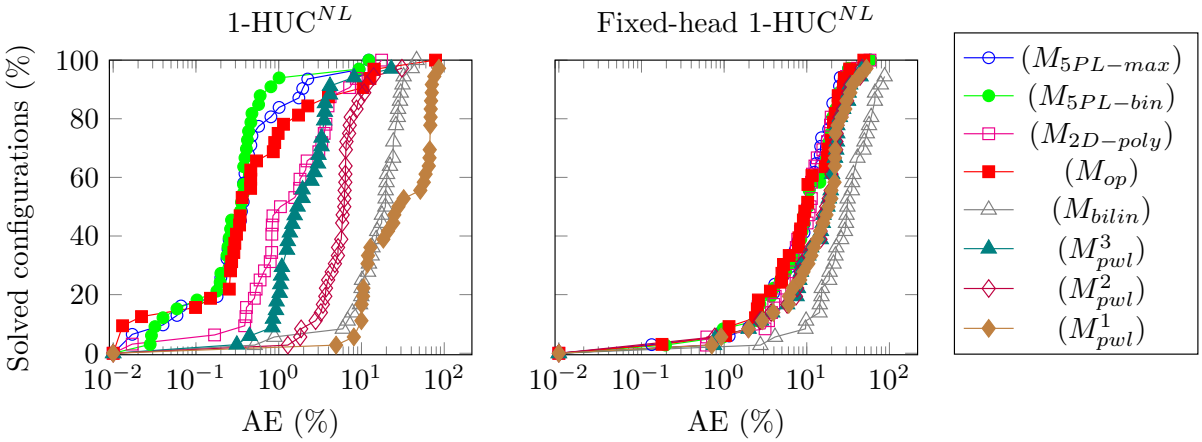


Figure 11: Proportion of configurations solved with their VBS where the AE is under an AE threshold

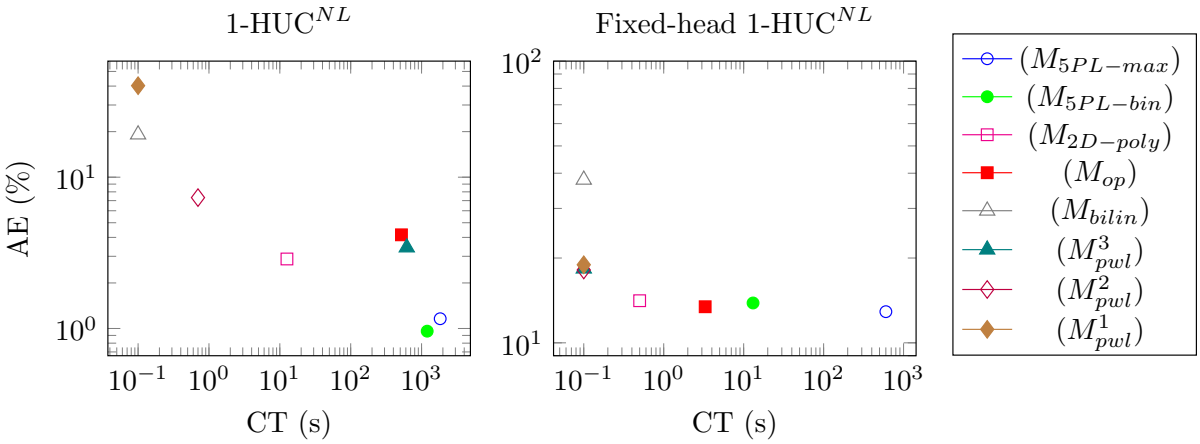


Figure 12: Trade-off between the average CT and AE for each model

and (P_{bilin}) yield a DB above 1% for nearly all instances. Besides, models with the smallest DB tend to be the ones with smaller AE. The only exception is model (P_{op}) with high DB despite a low AE.

The reason why the DB is correlated to the AE is because models with lower AE correspond more precisely to the physics. Hence, the solutions obtained with these models are closer to the real optimal solution of (P_{ref}) than the solutions obtained with models with a higher AE. Model (P_{op}) is an exception. Indeed, model (P_{op}) has a similar AE as $(P_{5PL-max})$ and $(P_{5PL-bin})$ as shown in **Figure 11**. However the DB for (P_{op}) is much higher than for $(P_{5PL-max})$ and $(P_{5PL-bin})$ (**Figure 13**), which means that its solutions are of lesser economic

quality. This is because for (P_{op}), there is a finite set of operating flows, and every solution must have water flows within this set. It is possible that the optimal solutions obtained with (P_{op}) are far from the optimal solutions of (P_{ref}) for which the water flows are not restricted to a finite set.

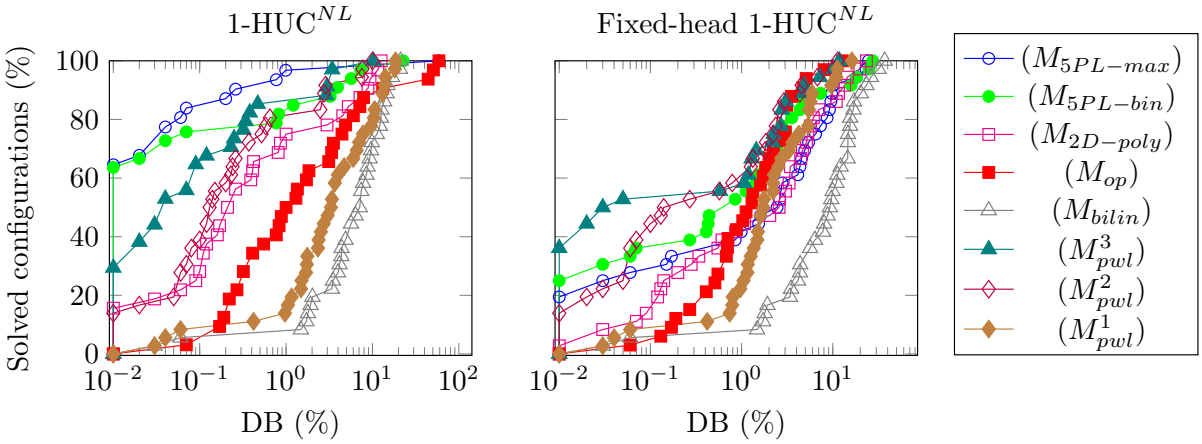


Figure 13: Proportion of configurations solved with their VBS where the DB is under a DB threshold

Impact of each characteristic. We now summarize the impact of modifying the characteristics of the 1-HUC problem. In **Table 7**, we present the general impact for the 1-HUC problem when modifying one characteristic. One arrow up (resp. down) means a moderate increase (resp. decrease), and two arrows up (resp. down) means a large increase (resp. decrease). We also added remarks for some models. The results presented here are further described in **Appendix D**, and the corresponding tables are shown in **Appendix E**.

Table 7: Summary of the impact of each characteristic

modified characteristics	1-HUC	fixed-head 1-HUC	remarks
increased size	CT↑↑	AE↑↑	CT increases much more for slower models
added equality constraints	CT↓, AE↓↓	AE↓↓	(P_{op}) can yield infeasibilities
increased degree of non-linearity	AE↑↑	AE↑↑	$(P_{2D-poly})$ can yield infeasibilities
increased number of inflection points	CT↑	CT↑	CT increases only for $(P_{2D-poly})$, $(P_{5PL-bin})$ and $(P_{5PL-max})$
decreased sensitivity to decision variables	CT↓	AE↓↓	all variants of (P_{pwl}) can yield high AE

5.4. Solvers comparison

Previous results are presented with respect to the *virtual best solver* (VBS). However, in a practical case it may not be convenient to use the VBS, as it could be difficult to have access to as many solvers. In this section we will analyze the behaviour of each solver independently. As aforementioned, a solved configuration has, for every solver, the same AE. Only the proportion of configurations solved and the CT can change from a solver to another. Hence, we do

not consider metric AE when comparing the solvers. In the tables presented in this section, the following notations are used:

- %S: proportion of configurations solved;
- avg-CT: average CT in seconds, for solved configurations;
- NS: model not supported by the solver;
- NR: model supported by the solver, but experiments are not reported.

For a model, a solver dominates another solver if it has a higher %S, and a smaller avg-CT.

Finding 1: The performance of a solver highly depends on the model. **Table 8** and **Table 9** show for each model the proportion of configurations solved with each solver, and the average CT, respectively for the 1-HUC problem and for the fixed-head 1-HUC problem. For each model, the smallest avg-CT and the highest %S are emphasized in bold. If for a model, both metrics are in bold for a solver, then it dominates every other solver for the model.

The results confirm that the performance of a solver highly depends on the model. For instance, when the 1-HUC problem is considered (**Table 8**), solver SCIP dominates solver ANTIGONE for model ($P_{5PL-bin}$), but ANTIGONE dominates SCIP for model ($P_{2D-poly}$).

Table 8: Proportion of configurations for each model solved by each solver and related average CT for the 1-HUC problem

	ANTIGONE		BARON		COUENNE		LINDOGlobal		SCIP		CPLEX	
Model	%S	avg-CT	%S	avg-CT	%S	avg-CT	%S	avg-CT	%S	avg-CT	%S	avg-CT
($P_{5PL-max}$)	NS		NS		NS		38.9	616.36	**86.1**	**606.87**	NS	
($P_{5PL-bin}$)	19.4	824.84	**88.9**	**25.63**	69.4	163.23	86.1	446.92	**88.9**	218.00	NS	
($P_{2D-poly}$)	44.4	**2.51**	**88.9**	31.32	63.9	91.27	80.6	30.36	19.4	828.72	NS	
(P_{op})	50.0	235.24	**83.3**	258.94	52.8	439.83	75.0	375.10	69.4	**44.23**	NS	
(P_{bilin})	75.0	0.10	**100.0**	**0.08**	97.2	0.25	75.0	4.72	**100.0**	0.17	NS	
(P^3_{pwl})	NR		NR		NR		NR		NR		**94.4**	**19.07**
(P^2_{pwl})	NR		NR		NR		NR		NR		**100.0**	**0.71**
(P^1_{pwl})	NR		NR		NR		NR		NR		**100.0**	**0.02**

Table 9: Proportion of configurations for each model solved by each solver and related average CT for the fixed-head 1-HUC problem

	ANTIGONE		BARON		COUENNE		LINDOGlobal		SCIP		CPLEX	
Model	%S	avg-CT	%S	avg-CT	%S	avg-CT	%S	avg-CT	%S	avg-CT	%S	avg-CT
($P_{5PL-max}$)	NS		NS		NS		**97.2**	159.96	94.4	**0.13**	NS	
($P_{5PL-bin}$)	30.6	856.07	**100.0**	13.13	**100.0**	109.84	61.1	58.80	**100.0**	**2.95**	NS	
($P_{2D-poly}$)	72.2	**0.15**	**100.0**	0.51	100.0	0.87	**100.0**	1.36	**100.0**	100.31	NS	
(P_{op})	NR		NR		NR		NR		NR		**91.7**	**0.01**
(P_{bilin})	NR		NR		NR		NR		NR		**100.0**	**0.00**
(P^3_{pwl})	NR		NR		NR		NR		NR		**100.0**	**0.05**
(P^2_{pwl})	NR		NR		NR		NR		NR		**100.0**	**0.02**
(P^1_{pwl})	NR		NR		NR		NR		NR		**100.0**	**0.01**

Finding 2: Solvers SCIP , BARON and CPLEX are the most efficient ones. **Tables 10** and **11** show the proportion of instances where a solver is the VBS, for each model.

Solver SCIP is the most efficient for model ($P_{5PL-max}$). Solver BARON is the most efficient for any other non-linear model, namely models ($P_{5PL-bin}$), ($P_{2D-poly}$), (P_{bilin}) and (P_{op}) for the 1-HUC problem and models ($P_{5PL-bin}$) and ($P_{2D-poly}$) for the fixed-head 1-HUC problem. Solver CPLEX is efficient for any linear model, namely models (P^1_{pwl}), (P^2_{pwl}) and (P^3_{pwl}) for the 1-HUC problem and models (P_{bilin}), (P_{op}) for the fixed-head 1-HUC problem.

We distinguish two exceptions where solver ANTIGONE is the most efficient: for model (P_{bilin}) for the 1-HUC problem and model ($P_{2D-poly}$) for the fixed-head 1-HUC problem. However, **Tables 8** and **9** show that in these cases, the problem is solved very quickly. Thus it is possible that ANTIGONE is quicker than BARON only on easy configurations, so it may only be quicker to start-up.

Table 10: Proportion of configurations for each model where a solver is the VBS for the 1-HUC problem

Model	ANTIGONE	BARON	COUENNE	LINDOGlobal	SCIP	CPLEX
($P_{5PL-max}$)	NS	NS	NS	16.1	**83.9**	NS
($P_{5PL-bin}$)	0.0	**54.6**	30.3	15.2	0.0	NS
($P_{2D-poly}$)	12.5	**84.4**	3.1	0.0	0.0	NS
(P_{op})	0.0	**90.6**	0.0	6.3	3.1	NS
(P_{bilin})	**63.9**	33.3	0.0	0.0	2.8	NS
(P^3_{pwl})	NR	NR	NR	NR	NR	**100.0**
(P^2_{pwl})	NR	NR	NR	NR	NR	**100.0**
(P^1_{pwl})	NR	NR	NR	NR	NR	**100.0**

Table 11: Proportion of configurations for each model where a solver is the VBS for the fixed-head 1-HUC problem

Model	ANTIGONE	BARON	COUENNE	LINDOGlobal	SCIP	CPLEX
($P_{5PL-max}$)	NS	NS	NS	5.6	**94.4**	NS
($P_{5PL-bin}$)	2.8	**72.2**	0.0	13.9	11.1	NS
($P_{2D-poly}$)	**52.8**	33.3	0.0	13.9	0.0	NS
(P_{op})	NR	NR	NR	NR	NR	**100.0**
(P_{bilin})	NR	NR	NR	NR	NR	**100.0**
(P^3_{pwl})	NR	NR	NR	NR	NR	**100.0**
(P^2_{pwl})	NR	NR	NR	NR	NR	**100.0**
(P^1_{pwl})	NR	NR	NR	NR	NR	**100.0**

In **Appendix D** an analysis of the impact of each feature of a 1-HUC problem instance is described.

5.5. General modeling recommendations derived from the numerical experiments

The results show that the choice of the model highly depends on the needs of a user. Indeed, if a low AE is required, models ($P_{5PL-bin}$), and ($P_{5PL-max}$) are the most efficient. However, not all solvers support ($P_{5PL-max}$), and larger

CT can be induced, meaning that ($P_{5PL-bin}$) is overall a better alternative than ($P_{5PL-max}$). For the fixed-head 1-HUC problem, models (P_{op}) and ($P_{2D-poly}$) can also be considered, as they yield nearly the same AE as ($P_{5PL-bin}$), and ($P_{5PL-max}$). In opposition, if one requires low computational times, models (P_{bilin}), and (P^1_{pwl}) are the most suitable. However, the AE can be very high with such models. For a more balanced option, three types of models stand out: (P_{pwl}) and ($P_{2D-poly}$) for the 1-HUC problem and (P_{pwl}), (P_{op}) and ($P_{2D-poly}$) for the fixed-head 1-HUC problem. We discuss them hereafter, giving for each of them their main strengths and weaknesses. Firstly, (P_{pwl}) usually provides a good trade-off between CT and AE. However the proper number of pieces cannot always be deduced in advance. Consequently, a trial-and-error procedure may be necessary to determine a piecewise linear function with a good trade-off. Secondly, model (P_{op}) can lead to the smallest AE, and it can be solved faster than the sophisticated models ($P_{5PL-bin}$) and ($P_{5PL-max}$). The drawback is that in the case of an instance with equality constraints there may not be a feasible solution for model (P_{op}). Thirdly, model ($P_{2D-poly}$) also yields a very good trade-off between AE and CT. However, for some configurations featuring model ($P_{2D-poly}$), all solvers fail to find a solution, even if there is a feasible solution with the model. This illustrates the intrinsic difficulties of the current solvers for some non-linear models (see also the case of ($P_{HD-poly}$) described in **Section 5.3**).

The choice of the solver impacts the CT and the proportion of instances solved. The results indicate that BARON is the most efficient non-linear solver when the model is supported, otherwise SCIP is the most efficient one. For the three balanced models highlighted, solver BARON is the most efficient for the non-linear ones (model (P_{op}) for the 1-HUC problem and ($P_{2D-poly}$) in any case), and a specialized MILP solver should be considered for the linear ones (model (P_{op}) for the fixed-head 1-HUC problem and (P_{pwl}) in any case).

6. Conclusion

In this paper various non-linear and linear modeling alternatives to solve a non-linear problem are compared, in terms of feasibility, approximation error, distance to the best recalculated value and computational time. The considered non-linear problem is the 1-HUC, featuring two non-linearities: a one-dimensional concave function, and a two-dimensional non-convex and non-concave function. A common special case of the 1-HUC, the fixed-head 1-HUC, is also considered, featuring a single non-linearity: a one-dimensional non-convex and non-concave function. A simplified model, with the power as a non-linear function of both the turbing efficiency and the head is defined for the 1-HUC and for the fixed-head 1-HUC. However, this model features too difficult non-linearities to be solved in a reasonable time, even for small instances. Seven alternative models are proposed, the focus being to represent the non-linearities of both the 1-HUC and the fixed-head 1-HUC. These models cover a large panel of modeling alternatives, including the common models for the 1-HUC from the literature, but also new models with uncommon non-linear functions. Several

sets of instances with different features of the 1-HUC and the fixed-head 1-HUC are solved with each of the proposed model, using five global solvers, and one linear solver. The results show that three of the seven models, namely (P_{op}), ($P_{2D-poly}$) and (P_{pwl}), stand out as the most appealing, offering the best trade-off between computational time, approximation error and feasibility. As the computing time of a non-linear model highly depends on the available global solver, preferred solvers are also highlighted for these three models.

As future research, refining the three most efficient models revealed by the present study, via advanced or dedicated solution methods, is promising. First, one can reformulate the models, using logarithmic disjunctive constraints [51]. This would benefit to all models, as it would reduce the number of variables. Second, introducing bound-tightening techniques can also lead to better results. Indeed we relied on the bound-tightening provided by the solvers, but one might induce tighter bounds dedicated to the 1-HUC. Third, model (P_{pwl}) can be improved by extending methods from the literature that optimize the number of breakpoints with an approximation guarantee in a PWL bounding framework [34]. Fourth, as model ($P_{2D-poly}$) features quadratic constraints, making use of quadratic programming techniques could improve the feasibility and the computational time. Finally, in opposition to the other presented models, the water flow is discretized in model (P_{op}) due to the finite set of operating points. Consequently there is a large combinatorics, which usually increases exponentially with the size of the instances. The use of combinatorial optimization methods, such as a polyhedral study could lead to smaller computing times. One interesting approach for future work would be to solve the 1-HUC with a fast model in order to quickly provide an initial solution to a more precise model. If the solution provided is better than the first solution obtained by the solvers, this approach could greatly reduce the computational time, yet still leading to a low approximation error. Lastly, MILP commercial solvers such as CPLEX and Gurobi [26] are currently being extended in order to solve non-linear problems. At the moment, very few of our models are supported, and the optimum is not guaranteed. However, these solvers might take into account a larger set of our models in the future, and may be more efficient than the non-linear solvers considered in the paper.

References

[1] Amani, A. and Alizadeh, H. (2021). Solving hydropower unit commitment problem using a novel sequential mixed integer linear programming approach. *Water Resources Management*, 35:1711–1729.

[2] Androulakis, I. P. (2001). MINLP: Branch and bound global optimization algorithm. *Encyclopedia of Optimization*, pages 1415–1421.

[3] Arce, A., Ohishi, T., and Soares, S. (2002). Optimal dispatch of generating units of the itaipu hydroelectric plant. *IEEE Transactions on power systems*, 17(1):154–158.

[4] Bellman, R. (1966). Dynamic programming. *Science*, 153(3731):34–37.

[5] Belotti, P. (2009). Couenne: a user's manual. Technical report, Lehigh University.

[6] Borghetti, A., D'Ambrosio, C., Lodi, A., and Martello, S. (2008). An MILP approach for short-term hydro scheduling and unit commitment with head-dependent reservoir. *IEEE Transactions on power systems*, 23(3):1115–1124.

[7] Brandao, L. C., Pessanha, J. F., Khenayfis, L. S., Diniz, A. L., Pereira, R. J. C., and Junior, C. A. A. (2022). A data-driven representation of aggregate efficiency curves of hydro units for the mid-term hydrothermal coordination problem. *Electric Power Systems Research*, 212:108511.

[8] Catalão, J. P. d. S., Pousinho, H. M. I., and Mendes, V. M. F. (2010). Mixed-integer nonlinear approach for the optimal scheduling of a head-dependent hydro chain. *Electric Power Systems Research*, 80(8):935–942.

[9] Conejo, A. J., Arroyo, J. M., Contreras, J., and Villamor, F. A. (2002). Self-scheduling of a hydro producer in a pool-based electricity market. *IEEE Transactions on power systems*, 17(4):1265–1272.

[10] Cplex, I. I. (2009). V12. 1: User's manual for CPLEX. *International Business Machines Corporation*, 46(53):157.

[11] Croxton, K. L., Gendron, B., and Magnanti, T. L. (2003). A comparison of mixed-integer programming models for nonconvex piecewise linear cost minimization problems. *Management Science*, 49(9):1268–1273.

[12] Czyzyk, J., Mesnier, M. P., and Moré, J. J. (1998). The NEOS server. *IEEE Computational Science and Engineering*, 5(3):68–75.

[13] Desrosiers, J. and Lübbecke, M. E. (2011). Branch-price-and-cut algorithms. *Encyclopedia of Operations Research and Management Science. John Wiley & Sons, Chichester*, pages 109–131.

[14] D'Ambrosio, C., Lodi, A., and Martello, S. (2010). Piecewise linear approximation of functions of two variables in MILP models. *Operations Research Letters*, 38(1):39–46.

[15] e Souza, H. G., Finardi, E. C., Brito, B. H., and Takigawa, F. Y. K. (2022). Partitioning approach based on convex hull and multiple choice for solving hydro unit-commitment problems. *Electric Power Systems Research*, 211:108285.

[16] Edom, E., Anjos, M., D'Ambrosio, C., Van Ackooij, W., Côté, P., and Séguin, S. (2020). On the impact of the power production function approximation on hydropower maintenance scheduling. *Les Cahiers du GERAD G–2020–22, GERAD*.

[17] Epperly, T. G. and Pistikopoulos, E. N. (1997). A reduced space branch and bound algorithm for global optimization. *Journal of Global Optimization*, 11(3):287–311.

[18] Finardi, E., Takigawa, F., and Brito, B. (2016). Assessing solution quality and computational performance in the hydro unit commitment problem considering different mathematical programming approaches. *Electric Power Systems Research*, 136:212–222.

[19] Finardi, E. C. and da Silva, E. L. (2006). Solving the hydro unit commitment problem via dual decomposition and sequential quadratic programming. *IEEE transactions on Power Systems*, 21(2):835–844.

[20] Finardi, E. C. and Scuzziato, M. R. (2013). Hydro unit commitment and loading problem for day-ahead operation planning problem. *International Journal of Electrical Power & Energy Systems*, 44(1):7–16.

[21] García-González, J., Parrilla, E., Barquín, J., Alonso, J., Sáiz-Chicharro, A., and González, A. (2003). Under-relaxed iterative procedure for feasible short-term scheduling of a hydro chain. In *2003 IEEE Bologna Power Tech Conference Proceedings,*, volume 2, page 6 pp.

[22] Geißler, B., Martin, A., Morsi, A., and Schewe, L. (2012). Using piecewise linear functions for solving MINLPs. In *Mixed integer nonlinear programming*, volume 154 of *The IMA Volumes in Mathematics and its Applications*, pages 287–314. Springer.

[23] Glasnovic, Z. and Margeta, J. (2009). The features of sustainable solar hydroelectric power plant. *Renewable energy*, 34(7):1742–1751.

[24] Gottschalk, P. G. and Dunn, J. R. (2005). The five-parameter logistic: a characterization and comparison with the four-parameter logistic. *Analytical biochemistry*, 343(1):54–65.

[25] Guignard, M. and Kim, S. (1987). Lagrangean decomposition: A model yielding stronger lagrangean bounds. *Mathematical programming*, 39(2):215–228.

[26] Gurobi Optimization, LLC (2023). Gurobi Optimizer Reference Manual.

[27] Kerschke, P., Kotthoff, L., Bossek, J., Hoos, H. H., and Trautmann, H. (2018). Leveraging TSP solver complementarity through machine learning. *Evolutionary computation*, 26(4):597–620.

[28] Land, A. H. and Doig, A. G. (2010). An automatic method for solving discrete programming problems. In *50 Years of Integer Programming 1958-2008*, pages 105–132. Springer.

[29] Li, X., Li, T., Wei, J., Wang, G., and Yeh, W. W.-G. (2013). Hydro unit commitment via mixed integer linear programming: A case study of the three gorges project, china. *IEEE Transactions on Power Systems*, 29(3):1232–1241.

[30] Lima, R. M., Marcovecchio, M. G., Novais, A. Q., and Grossmann, I. E. (2013). On the computational studies of deterministic global optimization of head dependent short-term hydro scheduling. *IEEE Transactions on Power Systems*, 28(4):4336–4347.

[31] Lin, Y. and Schrage, L. (2009). The global solver in the LINDO api. *Optimization Methods & Software*, 24(4-5):657–668.

[32] Mariano, S., Catalão, J., Mendes, V., and Ferreira, L. (2008). Optimising power generation efficiency for head-sensitive cascaded reservoirs in a competitive electricity market. *International Journal of Electrical Power & Energy Systems*, 30(2):125–133.

[33] Misener, R. and Floudas, C. A. (2014). ANTIGONE: algorithms for continuous/integer global optimization of nonlinear equations. *Journal of Global Optimization*, 59(2-3):503–526.

[34] Ngueveu, S. U. (2019). Piecewise linear bounding of univariate nonlinear functions and resulting mixed integer linear programming-based solution methods. *European Journal of Operational Research*, 275(3):1058–1071.

[35] Orero, S. and Irving, M. (1998). A genetic algorithm modelling framework and solution technique for short term optimal hydrothermal scheduling. *IEEE Transactions on Power Systems*, 13(2):501–518.

[36] Padberg, M. and Rinaldi, G. (1991). A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. *SIAM review*, 33(1):60–100.

[37] Paredes, M., Martins, L., and Soares, S. (2014). Using semidefinite relaxation to solve the day-ahead hydro unit commitment problem. *IEEE Transactions on Power Systems*, 30(5):2695–2705.

[38] Pearl, J. (1984). *Heuristics: intelligent search strategies for computer problem solving*. Addison-Wesley Longman Publishing Co., Inc.

[39] Pérez, J. I. and Wilhelmi, J. R. (2007). Nonlinear self-scheduling of a single unit small hydro plant in the day-ahead electricity market. *Proc. of ICREPQ'07*.

[40] Quesada, I. and Grossmann, I. E. (1995). Global optimization of bilinear process networks with multicomponent flows. *Computers & Chemical Engineering*, 19(12):1219–1242.

[41] Renaud, A. (1993). Daily generation management at Electricité de France: from planning towards real time. *IEEE Transactions on Automatic Control*, 38(7):1080–1093.

[42] Rodriguez, J. A., Anjos, M. F., Côté, P., and Desaulniers, G. (2018). Milp formulations for generator maintenance scheduling in hydropower systems. *IEEE Transactions on Power Systems*, 33(6):6171–6180.

[43] Ryoo, H. S. and Sahinidis, N. V. (1996). A branch-and-reduce approach to global optimization. *Journal of global optimization*, 8(2):107–138.

[44] Sahraoui, Y., Bendotti, P., and d'Ambrosio, C. (2019). Real-world hydropower unit-commitment: Dealing with numerical errors and feasibility issues. *Energy*, 184:91–104.

[45] Savelsbergh, M. (1997). A branch-and-price algorithm for the generalized assignment problem. *Operations research*, 45(6):831–841.

[46] Séguin, S., Côté, P., and Audet, C. (2015). Self-scheduling short-term unit commitment and loading problem. *IEEE Transactions on Power Systems*, 31(1):133–142.

[47] Skjelbred, H. I., Kong, J., and Fosso, O. B. (2020). Dynamic incorporation of nonlinearity into MILP formulation for short-term hydro scheduling. *International journal of electrical power & energy systems*, 116:105530.

[48] Smith, E. M. and Pantelides, C. C. (1999). A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering*, 23(4-5):457–478.

[49] Tawarmalani, M. and Sahinidis, N. V. (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103:225–249.

[50] Vaidyanathan, R. and El-Halwagi, M. (1996). Global optimization of nonconvex MINLP's by interval analysis. In *Global optimization in engineering design*, pages 175–193. Springer.

[51] Vielma, J. P. and Nemhauser, G. L. (2011). Modeling disjunctive constraints with a logarithmic number of binary variables and constraints. *Mathematical Programming*, 128(1-2):49–72.

[52] Vigerske, S. and Gleixner, A. (2018). SCIP: Global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. *Optimization Methods and Software*, 33(3):563–593.

[53] Zamora, J. M. and Grossmann, I. E. (1999). A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms. *Journal of Global Optimization*, 14(3):217–249.

Appendices

A. Solver description

ANTIGONE [33] is based on an sBB algorithms. The problem is reformulated in order to find special structures. Once the structures are found, the relaxation of the problem is solved. The search space is split and the process repeated until convergence of the upper and the lower bounds. Upper bounds are computed with local optimization algorithms. Only twice differentiable functions, that are not trigonometrical functions, are supported by ANTIGONE.

BARON [49] implements a deterministic Branch and Reduce algorithm. This algorithm contains constraint programming, interval analysis and duality techniques for tightening variables bounds. Heuristics, cutting planes and parallelism are combined with the Branch and Reduce algorithm. Trigonometrical functions and **max** functions are not supported.

COUENNE [5] implements an sBB with linearization, bound reductions and branching method. The main four components are: reformulation, separation of linearization cuts, branching rules and bound tightening methods. COUENNE only supports functions that can be reformulated into univariate functions and does not support function **max**.

LINDOGlobal [31] is the only solver that does not directly implements an sBB algorithm. Instead, it implements a branch and cut algorithm that breaks the model into sub-problems. The sub-problems are further split until each sub-problem is convex. The sub-problems are then solved with a BB or sBB algorithm. LINDOGlobal supports most non-linearities, and binary operators such as AND, OR and NOT.

SCIP [52] implements an sBB, where the non-linearities are represented within graphs. These graphs help finding convex non-linearities, and reformulating the non-linear functions. During the solving process, SCIP also adds various cuts, depending on the non-linearities. Bound tightening methods are also applied. Trigonometrical functions are not supported by SCIP and it is the only solver which requires a linear objective function.

CPLEX [10] implements a quite effective multipurpose Branch and Cut algorithm, which generates automatically various cuts [10]. Furthermore it is paired with pre-processing and heuristics.

B. Five parameters logistic function

A $\mathbf{5PL}$ is the following function, where x is a variable and y_1 to y_5 the parameters:

$$5PL(x, y_1, y_2, y_3, y_4, y_5) = y_4 + \frac{-y_4}{(1 + (\frac{x-y_1}{y_3})^{y_2})^{y_5}}$$

In the context of the 1-HUC, variable x is the water-flow d_t. The $\mathbf{5PL}$ has a shape similar to a more common function, the sigmoid:

$$sig(x, y'_1, y'_2, y'_3) = \frac{y'_3}{1 + e^{-y'_2(x-y'_1)}}$$

The advantages of the $\mathbf{5PL}$ is that it is more flexible than a sigmoid. The sigmoid is necessarily symmetric with respect to its inflection point, whereas $\mathbf{5PL}$ is not. However, a $\mathbf{5PL}$ function is not defined if $x < y_1$, which can occur when representing a unit by a $\mathbf{5PL}$ function. To adapt the $\mathbf{5PL}$ function to the use case of the 1-HUC, it is possible to insert a **max** function inside the $\mathbf{5PL}$ function as follows:

$$5PL(x, y_1, y_2, y_3, y_4, y_5) = y_4 + \frac{-y_4}{\left(1 + \left(\frac{\mathbf{max}(0, x-y_1)}{y_3}\right)^{y_2}\right)^{y_5}}$$

With this modification, if $x < y_1$ then the $\mathbf{5PL}$ function is equal to $y_4 + (-y_4/1) = 0$, if $x \geq y_1$, the $\mathbf{5PL}$ has the same behaviour as previously defined.

C. Instances description

The instances are derived from the following parameter sets A and B, by changing the value of only one parameter at a time. The idea is to evaluate the impact of the parameters on the resolution and the solution with multiple metrics. **Table 12** shows the parameters of each parameter set.

Table 12: Parameter sets

Parameter set A	Parameter set B
$V_0^1 = 500$, $V_0^2 = 200$	$V_0^1 = 90$, $V_0^2 = 10$
$T = 4$	$T = 4$
$\overline{V}_t^1 = 1000$, $\underline{V}_t^1 = 0\ \forall t \leq T$	$\overline{V}_t^1 = 100$, $\underline{V}_t^1 = 0\ \forall t \leq T$
$\overline{V}_t^2 = 500$, $\underline{V}_t^2 = 0\ \forall t \leq T$	$\overline{V}_t^2 = 90$, $\underline{V}_t^2 = 0\ \forall t \leq T$
$\underline{D} = 0$, $\overline{D} = 25$	$\underline{D} = 0$, $\overline{D} = 8$
$\underline{P}_t = 0$, $\overline{P}_t = 15\ \forall t \leq T$	$\underline{P}_t = 0$, $\overline{P}_t = 32\ \forall t \leq T$
$\Phi^1 = 230$, $\Phi^2 = 0$	$\Phi^1 = 850$, $\Phi^2 = 0$
$\Lambda = [0.2, 0.15, 0.1, 0.2]$	$\Lambda = [0.1, 0.2, 0.5, 0.4]$
$A_t^1 = A_t^2 = 0\ \forall t \leq T$	$A_t^1 = A_t^2 = 0\ \forall t \leq T$
$\gamma = [0, 0.1, 5, 0.7]$	$\gamma = [100, 0.2, 2, 0.6]$

The modified features are the following:

- Size of the instance;
- Equality constraints;

- Number of inflection points of the non-linear function;
- Degree of non-linearity of the function;
- Sensitivity of the decision variables to the non-linear effect.

In these parameter sets, the maximum and minimum volumes are artificially large, to see the impact of each feature. Below, we justify the choice for each feature and explain how the changes are instantiated on the 1-HUC. Note that every instance is built such that there is at least one feasible solution with continuous water flows d_t.

C.1. Size of the instance

Larger instances are in general harder to solve as they contain more variables and constraints. In the case of the 1-HUC, larger instances considered will have a larger number of time periods T. Increasing T exponentially increases the number of feasible solutions. Three instances are considered, A-T-1 to A-T-3 (resp. B-T-1 to B-T-3) corresponding to the variations of the parameter set A (resp. B) with 4, 7 and 10 time periods T. To take into account more time periods, prices are supplemented as follows: Λ = [0.2, 0.15, 0.1, 0.2, 0.1, 0.05, 0.1, 0.2, 0.15, 0.05] (resp. Λ = [0.1, 0.2, 0.5, 0.4, 0.3, 0.2, 0.3, 0.5, 0.4, 0.2]). These instances are such that the volume of each reservoir can not reach the maximum or minimum volume. The water flow will not be affected by the bounds on the volume, in contrary to some other sets of instances.

C.2. Equality constraints

Equality constraints can highly affect the resolution. Indeed, equality constraints drastically reduce the number of feasible solutions and can also be hard to satisfy. Moreover, depending on the approximation used in the model, equality constraints may lead to non efficient solutions. In the case of the 1-HUC, target volumes are equality constraints, when $\underline{V}_t^1 = \overline{V}_t^1$ for a time period t. Six instances are considered, A-E-1 to A-E-6 (resp. B-E-1 to B-E-6) which are variations of parameter set A (resp. B), where target volumes are only for the last time period T. For A-E-1 to A-E-3 (resp. B-E-1 to B-E-3) the target volumes are 480, 450 and 420 (resp. 80, 70 and 60). For A-E-4 to A-E-6 (resp B-E-4 to B-E-6), the target volumes are 500 (resp. 90), but the additional intake of water at the last time period are 20, 50 and 80 (resp. 10, 20, 30). One can notice that for instance A-E-1, the difference between the initial and the target volume is 20, while for instance A-E-4 it is 0, but the additional intake of water is 20. Thus, feasible solutions for A-E-1 are feasible solutions for A-E-4 and vice-versa. Instances A-E-2 and A-E-5, B-E-1 and B-E-4 and so on are built similarly.

C.3. Number of inflection points of the non-linear function

With a different number of inflection points, the shape of a non-linear function is changed, which can lead to more local optimal solutions, or less efficient under-estimators. The functions used to under-estimate and approximate the functions are also changed. Thus, the resolution and approximation error could be impacted by the number of inflection points of the non-linear function. In the case of the 1-HUC, the number of inflection points of the power function can be changed by defining a larger number of smaller units. We still have the same maximum power and maximum water flow, only the shape of the power function is different. Three instances are considered, A-N-1 to A-N-3 (resp. B-N-1 to B-N-3) which are variations of the parameter set A (resp. B) with 2, 4 and 6 units.

C.4. Degree of non-linearity of the non-linear function

As for the number of inflection points, changing the degree of non-linearity is another way to change the shape of a function. Thus, the resolution and approximation error could be affected by the degree of non-linearity of the functions. In the case of the 1-HUC, one way to increase the non-linearity of the power function is to change the water flow when each unit starts and stops, increasing or reducing the degree of curvature for each concave part of the function, as represented by. The resulting power function for the plant can be quasi-linear or have a high degree of non-linearity. In addition, it also changes the domain of some variables. Six instances are considered, A-D-1 to A-D-6 (resp. B-D-1 to B-D-6) being variations of parameter set A (resp. B). Instances A-D-1 and A-D-4 feature a quasi linear function, with $\overline{D} = 22$ and $\overline{P}_t = 14.5$, $\forall t \leq T$, instances A-D-2 and A-D-5 correspond to a non-linear function, with $\overline{D} = 25$ and $\overline{P}_t = 15$, and instances A-D-3 and A-D-6 use a very non-linear function, with $\overline{D} = 28$ and $\overline{P}_t = 16$. The target volume for instances A-D-4 to A-D-6 is 460. Similarly, instances B-D-1 and B-D-4 feature a quasi linear function, with $\overline{D} = 6$ and $\overline{P}_t = 28$, instances B-D-2 and B-D-5 feature a non-linear function, with $\overline{D} = 8$ and $\overline{P}_t = 32$, and instances B-D-3 and B-D-6 feature a very non-linear function, with $\overline{D} = 10$ and $\overline{P}_t = 34$. The target volume for instances B-D-4 to B-D-6 is 75.

C.5. Sensitivity of the decision variables to the non-linear effect

Depending on the problem, decision variables can have a very large, or very small impact on the non-linearities. When the impact is small, it is possible that some simplifications of the problem would not induce large approximation errors. In the case of the 1-HUC, the sensitivity of the decision variables to the non-linear effect can change by considering larger or smaller reservoirs. Two instances are considered, A-S-1 and A-S-2 (resp. B-S-1 and B-S-2) are variations of parameter set A (resp. B). Instance A-S-2 is similar to A-S-1, but has all initial, maximal and minimal volumes multiplied by 100, and supplemented prices $\Lambda = [0.005, 0.00375, 0.0025, 0.005]$. Analogously, B-S-2 is similar to B-S-1 with initial, maximal and minimal volumes all multiplied by 100, and adapted

prices $\Lambda = [0.005, 0.01, 0.025, 0.02]$. The unitary prices Λ are reduced in order to obtain similar solutions for instances A-S-1 and A-S-2 (resp. B-S-1 and B-S-2).

One can compute bounds on the variation of the volume, by calculating the maximum and minimum water processed while respecting the capacities. These bounds give an interval for the final volume in the reservoirs. It is then possible to compute the maximum difference in terms of volume between two feasible solutions, and compare it to the capacity of the reservoirs in order to predict if the instance might induce high volume variations or not. The sensitivity S can be computed as follows:

$$S = \frac{\overline{D} \times T - \underline{D} \times T}{\min(\overline{V}_T^1 - \underline{V}_T^1, \overline{V}_T^2 - \underline{V}_T^2)}$$

For instance A-S-1, the sensitivity is $100/500 = 0.2$, for instance A-S-2: 0.002, for instance B-S-1: 0.36 and for instance B-S-2: 0.0036. Note that the parameter set A (resp. B) has the same sensitivity as instance A-S-1 (resp. B-S-1).

Table 9 summarizes these instances and their features.

Table 13: Instance features

Instances	Features	Modified parameter
A-T-1 to A-T-3 B-T-1 to B-T-3	Size of the instance	Number of time periods
A-E-1 to A-E-6 B-E-1 to B-E-6	Equality constraints	Different target volumes, with and without additional intakes of water
A-N-1 to A-N-3 B-N-1 to B-N-3	Number of inflection points	Number of units
A-D-1 to A-D-6 B-D-1 to B-D-6	Degree of non-linearity	Quasi-linear, non-linear or very linear function, with and without target volumes
A-S-1, A-S-2, B-S-1, B-S-2	Sensitivity of the decision variables	Size of the reservoirs

D. Analysis of the impact of the instance features

Let us analyse the impact of each feature of a 1-HUC instance on the resolution. The tables related to the results described in the following section are in **Appendix E**.

D.1. Size of the instance

Changing the number of time periods (instances A-T-1 to A-T-3 and B-T-1 to B-T-3) has a big effect on the resolution. Indeed, we see from **Table 14** and **Table 15** that configurations with more time periods require a drastically increased CT compared to configurations with fewer time periods. The most salient case is for the 1-HUC where with $T = 10$ the only configurations solved under three hours by their VBS are with one of the following four models: $(P_{2D-poly})$, (P_{bilin}), (P_{pwl}^1) and (P_{pwl}^2). Moreover, with $T = 7$ configurations with model $(P_{5PL-max})$ are never solved by their VBS within three hours. The

AE also increases for configurations with instances with more time periods. It is especially visible for the fixed-head 1-HUC, with $T = 10$ the minimal AE is around 20% using models ($P_{5PL-max}$), and the average AE are around 40% at least.

As there are more time periods, more variables and constraints are introduced, exponentially increasing the number of feasible solutions. The reason why the AE increases is due to the fact that errors are propagated through the time periods. Also, for the fixed-head 1-HUC, more time periods mean, in general, more water processed. The volume varies with a higher magnitude from the initial volume when there are more time periods, leading to larger AE when considering a fixed-head.

D.2. Equality constraints

Taking fixed target volumes (instances A-E-1 to A-E-6 , B-E-1 to B-E-6, A-D-1 to A-D-6 and B-D-1 to B-D-6) has a non-homogeneous impact on the resolution. From **Table 16** and **Table 17** we notice that configurations with target volumes reduce the CT required for the 1-HUC, compared to configurations without target volumes. For the AE, we notice multiple behaviours. For most models, configurations with target volumes yields to smaller average AE, but higher maximal AE, compared to configurations without target volumes. Non-represented results also showed that ANTIGONE solves less than 25% of configurations with target volumes whereas it solves more than 60% of configurations without target volumes. A similar but less marked behaviour is noticed for COUENNE.

The decreased CT is probably due to the fact that fewer solutions are feasible. The reduced AE are due to the target volume being very close to the initial volume for some instances. Less volume is processed, meaning a smaller power, and smaller AE. Besides, for the fixed-head 1-HUC, it also means less errors due to the fixed head, as the volume may not vary to much from the initial volume. We notice that some configurations with model (P_{op}) are not solved, for both the 1-HUC and the fixed-head 1-HUC. This is because the target volume may not be reachable with the finite set of water flows.

D.3. Degree of non-linearity

Changing the non-linearity of the power function (instances A-D-1 to A-D-6 and B-D-1 to B-D-6) can have an impact on the CT and the AE in the case of the 1-HUC, but only on the AE for the fixed-head 1-HUC. From **Table 18** and **Table 19**, we notice that configurations with pronounced non-linearities have larger CT for the 1-HUC than configurations with quasi-linear functions. The configurations with non-linear models also have larger AE with pronounced non-linear functions, for both the 1-HUC and the fixed-head 1-HUC. The AE for configurations with linear model is not affected. We also see that all the configurations with model ($P_{2D-poly}$) are infeasible with every solver when the instance has a pronounced non-linear function, even if there exist feasible solutions for the instance.

The general increase of the AE for non-linear models can be explained by two reasons. Firstly, by instance construction, the units are the same for every instances, and the water-flow interval for each unit is changed in order to have a different degree of non-linearity. As such, it is possible that fewer non-linear functions can closely approximate the function of (P_{ref}) on a larger interval. Secondly, a highly non-linear function can be harder to approximate by simpler functions, leading to larger AE for every model.

D.4. Number of inflection points

Changing the number of units (instances A-N-1 to A-N-3 and B-N-1 to B-N-3) has only a noticeable impact on models representing each unit explicitly, namely ($P_{5PL-max}$), ($P_{5PL-bin}$) and ($P_{2D-poly}$). Indeed, **Table 20** and **Table 21** show the increased CT required for configurations with these models and with instances with more units. Also, increased number of units reduces the degree of non-linearity. Thus it is possible to see similar behaviours as when changing the degree of non-linearity.

The reason why the CT increases for configurations with one of the four mentioned models and an instance with many units is because as they represent each unit explicitly, more variables and constraints are required.

D.5. Sensitivity of the decision variables to the non-linear effect

In order to have negligible variation of the volume, the volumes can be set to larger values than the water flows. **Table 22** and **Table 23** show that the CT tends to be smaller for configurations with large volumes compared to configurations with smaller volumes. Larger volumes usually lead to an improvement of the AE for the fixed-head 1-HUC. However, the maximal AE of PWL models can be very large with large volumes. More precisely, with a PWL models, half the configurations has a large AE, and the other half has a smaller AE, compared to configurations with small volumes

The improvement of AE for the fixed-head 1-HUC is because with small variations, the volume is very similar to the initial volume at any time period. The AE from the fixed-head becomes very small. The high AE of the PWL models can be explained as follows. These models only consider a family of univariate PWL function for a finite set of possible volumes. It is then possible that the volume is never similar to the volumes used by this family of functions.

E. Numerical experiments when partitioning instances

- %S: proportion of configuration solved;
- min-CT, max-CT, avg-CT: minimum, maximum, and average CT for every solved configurations;
- min-AE, max-AE, avg-AE: minimum, maximum and average AE for every solved configurations.

E.1. Size of the instance

Table 14 and **Table 15** represents the proportion of configurations with instances with 4, 7 and 10 time periods and each model solved by their VBS, and related minimum, maximum and average CT and AE.

Table 14: Proportion of configurations solved with their VBS, CT and AE statistics for the 1-HUC for different number of time periods (instances A-T-1 to A-T-3 and B-T-1 to B-T-3)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
T=4	$(P_{5PL-max})$	100.0	103.34	186.38	144.86	0.2	0.4	0.3
	$(P_{5PL-bin})$	100.0	15.28	28.47	21.88	0.2	0.4	0.3
	$(P_{2D-poly})$	100.0	0.43	0.69	0.56	0.8	3.9	2.4
	(P_{op})	100.0	1.47	2.25	1.86	0.3	0.4	0.3
	(P_{bilin})	100.0	0.03	0.05	0.04	24.3	26.1	25.2
	(P^3_{pwl})	100.0	0.4	7.06	3.73	1.3	3.5	2.4
	(P^2_{pwl})	100.0	0.09	0.18	0.14	5.8	6.5	6.2
	(P^1_{pwl})	100.0	0.01	0.02	0.01	11.8	67.9	39.9
T=7	$(P_{5PL-max})$	0	-	-	-	-	-	-
	$(P_{5PL-bin})$	100.0	10 367.97	10 367.97	10 367.97	0.4	0.4	0.4
	$(P_{2D-poly})$	100.0	8.31	36.18	22.25	0.8	3.3	2.0
	(P_{op})	100.0	78.69	574.13	326.41	4.0	14.3	9.2
	(P_{bilin})	100.0	0.06	0.26	0.16	24.1	31.3	27.7
	(P^3_{pwl})	100.0	39.12	39.12	39.12	1.3	1.3	1.3
	(P^2_{pwl})	100.0	0.15	2.88	1.51	5.7	8.7	7.2
	(P^1_{pwl})	100.0	0.03	0.06	0.04	27.9	67.9	47.9
T=10	$(P_{5PL-max})$	0	-	-	-	-	-	-
	$(P_{5PL-bin})$	0	-	-	-	-	-	-
	$(P_{2D-poly})$	100.0	44.71	344.08	194.39	0.8	14.1	7.5
	(P_{op})	100.0	7062.75	7062.75	7062.75	78.6	78.6	78.6
	(P_{bilin})	100.0	0.07	0.09	0.08	40.4	46.3	43.3
	(P^3_{pwl})	50.0	558.73	558.73	558.73	1.1	1.1	1.1
	(P^2_{pwl})	100.0	0.46	19.05	9.76	7.4	14.0	10.7
	(P^1_{pwl})	100.0	0.05	0.16	0.11	32.4	66.8	49.6

E.2. Equality constraints

Table 16 and **Table 17** represents the proportion of configurations with instances with and without target volumes and each model solved by their VBS, and related minimum, maximum and average CT and AE.

E.3. Degree of non-linearity

Table 18 and **Table 19** represents the proportion of configurations with instances with a quasi linear, a non-linear and a very non-linear function and each model solved by their VBS, and related minimum, maximum and average CT and AE.

E.4. Number of inflection points

Table 20 and **Table 21** represents the proportion of configurations with instances with 2, 6 and 6 units and each model solved by their VBS, and related minimum, maximum and average CT and AE.

Table 15: Proportion of configurations solved with their VBS, CT and AE statistics for the fixed-head 1-HUC for different number of time periods (instances A-T-1 to A-T-3 and B-T-1 to B-T-3)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
T=4	$(P_{5PL-max})$	100.0	0.11	0.12	0.11	10.7	20.6	15.7
	$(P_{5PL-bin})$	100.0	0.25	0.4	0.33	10.7	20.6	15.7
	$(P_{2D-poly})$	100.0	0.06	0.08	0.07	14.6	27.3	20.9
	(P_{op})	100.0	0.0	0.01	0.01	10.3	21.0	15.7
	(P_{bilin})	100.0	0.0	0	0.0	41.0	66.6	53.8
	(P^3_{pwl})	100.0	0.02	0.02	0.02	20.2	21.9	21.0
	(P^2_{pwl})	100.0	0.01	0.01	0.01	19.9	21.8	20.9
	(P^1_{pwl})	100.0	0.01	0.01	0.01	21.9	22.2	22.0
T=7	$(P_{5PL-max})$	100.0	0.11	0.14	0.12	12.4	32.1	22.2
	$(P_{5PL-bin})$	100.0	2.77	10.14	6.46	18.6	32.1	25.4
	$(P_{2D-poly})$	100.0	0.06	0.1	0.08	23.6	43.4	33.5
	(P_{op})	100.0	0.01	0.01	0.01	17.8	33.1	25.5
	(P_{bilin})	100.0	0.0	0	0.0	60.2	115.6	87.9
	(P^3_{pwl})	100.0	0.03	0.03	0.03	31.5	32.1	31.8
	(P^2_{pwl})	100.0	0.01	0.01	0.01	30.8	31.6	31.2
	(P^1_{pwl})	100.0	0.01	0.01	0.01	30.2	35.7	33.0
T=10	$(P_{5PL-max})$	100.0	0.11	0.15	0.13	19.6	59.2	39.4
	$(P_{5PL-bin})$	100.0	6.25	50.0	28.12	29.7	59.2	44.5
	$(P_{2D-poly})$	100.0	0.08	0.08	0.08	34.7	67.3	51.0
	(P_{op})	100.0	0.01	0.01	0.01	28.4	49.0	38.7
	(P_{bilin})	100.0	0.0	0	0.0	78.7	149.9	114.3
	(P^3_{pwl})	100.0	0.04	0.04	0.04	44.7	47.0	45.9
	(P^2_{pwl})	100.0	0.01	0.02	0.01	44.2	46.1	45.2
	(P^1_{pwl})	100.0	0.01	0.01	0.01	42.9	52.8	47.8

Table 16: Proportion of configurations solved with their VBS, CT and AE statistics for the 1-HUC with and without target volumes (instances A-T-1, B-T-1, A-E-1 to A-E-6, B-E-1 to B-E-6, A-D-1 to A-D-6 and B-D-1 to B-D-6)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
No target volume	$(P_{5PL-max})$	100.0	51.99	212.62	148.12	0.0	12.3	1.9
	$(P_{5PL-bin})$	100.0	4.08	31.84	21.97	0.0	12.3	1.9
	$(P_{2D-poly})$	100.0	0.4	0.81	0.53	0.0	7.2	2.8
	(P_{op})	100.0	1.47	2.95	1.94	0.1	12.7	1.9
	(P_{bilin})	100.0	0.03	0.06	0.04	19.4	32.3	26.2
	(P^3_{pwl})	100.0	0.4	8.98	3.6	0.9	4.2	2.5
	(P^2_{pwl})	100.0	0.09	0.18	0.13	4.4	6.5	5.8
	(P^1_{pwl})	100.0	0.01	0.03	0.02	10.0	70.6	39.8
target volume	$(P_{5PL-max})$	100.0	0.66	399.96	94.73	0.0	9.3	1.1
	$(P_{5PL-bin})$	100.0	0.1	18.31	7.66	0.0	9.3	0.8
	$(P_{2D-poly})$	100.0	0.09	0.48	0.29	0.2	4.2	1.4
	(P_{op})	100.0	0.56	23.96	3.81	0.0	10.5	1.2
	(P_{bilin})	100.0	0.03	0.13	0.07	0.5	27.4	12.4
	(P^3_{pwl})	100.0	0.14	1.0	0.44	0.8	22.9	3.4
	(P^2_{pwl})	100.0	0.01	0.17	0.07	1.3	30.8	6.6
	(P^1_{pwl})	100.0	0.01	0.02	0.01	4.9	69.4	33.0

E.5. Sensitivity of the decision variables to the non-linear effect

Table 22 and **Table 23** represents the proportion of configurations with instances with small and large volumes and each model solved by their VBS,

Table 17: Proportion of configurations solved with their VBS, CT and AE statistics for the fixed-head 1-HUC with and without target volumes (instances A-T-1, B-T-1, A-E-1 to A-E-6, B-E-1 to B-E-6, A-D-1 to A-D-6 and B-D-1 to B-D-6)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
No target volume	$(P_{5PL-max})$	100.0	0.1	0.12	0.11	7.3	24.1	15.9
	$(P_{5PL-bin})$	100.0	0.2	0.48	0.35	8.8	28.7	17.8
	$(P_{2D-poly})$	100.0	0.06	0.14	0.08	9.6	28.4	20.8
	(P_{op})	100.0	0.0	0.01	0.01	8.0	27.8	17.4
	(P_{bilin})	100.0	0.0	0	0.0	29.0	87.5	56.5
	(P^3_{pwl})	100.0	0.02	0.03	0.03	17.7	27.5	21.4
	(P^2_{pwl})	100.0	0.01	0.02	0.01	17.2	27.4	21.1
	(P^1_{pwl})	100.0	0.01	0.01	0.01	18.4	29.2	22.2
target volume	$(P_{5PL-max})$	100.0	0.1	0.93	0.2	2.0	19.5	7.6
	$(P_{5PL-bin})$	100.0	0.03	1.06	0.52	1.0	80.0	14.0
	$(P_{2D-poly})$	100.0	0.06	0.23	0.13	0.6	28.8	10.4
	(P_{op})	100.0	0.0	0.02	0.01	1.0	19.9	7.1
	(P_{bilin})	100.0	0.0	0	0.0	2.7	49.8	21.2
	(P^3_{pwl})	100.0	0.02	0.17	0.07	0.8	33.5	10.6
	(P^2_{pwl})	100.0	0.01	0.04	0.02	0.8	33.5	10.6
	(P^1_{pwl})	100.0	0.0	0.01	0.01	0.9	33.6	10.8

Table 18: Proportion of configurations solved with their VBS, CT and AE statistics for the 1-HUC for different degree of non-linearity (instances A-D-1 to A-D-6 and B-D-1 to B-D-6)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
Quasi linear	$(P_{5PL-max})$	100.0	2.41	399.96	162.15	0.0	0.5	0.3
	$(P_{5PL-bin})$	100.0	3.18	31.84	13.63	0.0	0.5	0.3
	$(P_{2D-poly})$	100.0	0.23	0.4	0.35	0.0	7.2	2.9
	(P_{op})	100.0	1.43	1.68	1.54	0.1	0.5	0.2
	(P_{bilin})	100.0	0.03	0.08	0.05	10.2	27.7	18.1
	(P^3_{pwl})	100.0	0.2	3.39	1.15	1.0	3.6	2.2
	(P^2_{pwl})	100.0	0.1	0.14	0.12	4.4	6.2	5.4
	(P^1_{pwl})	100.0	0.01	0.02	0.02	10.0	68.7	38.8
Non-linear	$(P_{5PL-max})$	100.0	43.21	239.32	148.84	0.2	0.4	0.3
	$(P_{5PL-bin})$	100.0	8.83	29.02	17.6	0.2	0.4	0.3
	$(P_{2D-poly})$	100.0	0.21	0.81	0.46	0.8	3.9	1.9
	(P_{op})	100.0	1.08	2.93	1.98	0.3	0.4	0.3
	(P_{bilin})	100.0	0.04	0.08	0.05	0.9	26.1	16.7
	(P^3_{pwl})	100.0	0.15	6.73	2.01	1.3	3.6	2.6
	(P^2_{pwl})	100.0	0.03	0.16	0.1	5.8	7.1	6.4
	(P^1_{pwl})	100.0	0.01	0.02	0.02	11.8	68.4	40.0
Very non-linear	$(P_{5PL-max})$	100.0	62.31	208.98	151.88	1.0	12.3	6.2
	$(P_{5PL-bin})$	100.0	0.56	27.81	15.87	0.0	12.3	5.7
	$(P_{2D-poly})$	0	-	-	-	-	-	-
	(P_{op})	100.0	0.83	3.58	2.29	0.9	12.7	6.3
	(P_{bilin})	100.0	0.03	0.1	0.06	7.8	32.3	21.7
	(P^3_{pwl})	100.0	0.37	8.98	2.87	0.9	4.2	2.6
	(P^2_{pwl})	100.0	0.09	0.17	0.14	5.2	7.2	6.3
	(P^1_{pwl})	100.0	0.02	0.03	0.02	10.0	70.6	40.0

and related minimum, maximum and average CT and AE.

Table 19: Proportion of configurations solved with their VBS, CT and AE statistics for the fixed-head 1-HUC for different degree of non-linearity (instances A-D-1 to A-D-6 and B-D-1 to B-D-6)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
Quasi linear	$(P_{5PL-max})$	100.0	0.1	0.13	0.11	7.1	18.4	10.2
	$(P_{5PL-bin})$	100.0	0.2	0.82	0.46	7.1	18.4	10.5
	$(P_{2D-poly})$	100.0	0.06	0.14	0.1	6.8	28.4	14.3
	(P_{op})	100.0	0.01	0.01	0.01	7.9	16.8	10.9
	(P_{bilin})	100.0	0.0	0	0.0	21.1	53.2	31.4
	(P^3_{pwl})	100.0	0.02	0.03	0.03	7.5	18.5	14.8
	(P^2_{pwl})	100.0	0.01	0.02	0.01	7.4	18.4	14.7
	(P^1_{pwl})	100.0	0.01	0.01	0.01	6.8	19.8	15.2
Non-linear	$(P_{5PL-max})$	100.0	0.11	0.17	0.12	6.2	20.6	11.7
	$(P_{5PL-bin})$	100.0	0.32	1.03	0.56	6.2	20.6	11.5
	$(P_{2D-poly})$	100.0	0.06	0.23	0.15	7.5	27.3	14.1
	(P_{op})	100.0	0.0	0.01	0.01	6.1	21.0	11.5
	(P_{bilin})	100.0	0.0	0	0.0	18.7	66.6	37.3
	(P^3_{pwl})	100.0	0.02	0.03	0.03	7.8	21.9	16.1
	(P^2_{pwl})	100.0	0.01	0.01	0.01	7.6	21.8	16.0
	(P^1_{pwl})	100.0	0.01	0.01	0.01	6.1	22.2	15.9
Very non-linear	$(P_{5PL-max})$	100.0	0.1	0.12	0.11	10.1	24.1	15.9
	$(P_{5PL-bin})$	100.0	0.35	1.01	0.56	10.3	28.7	19.9
	$(P_{2D-poly})$	100.0	0.06	0.19	0.11	6.8	26.8	15.7
	(P_{op})	100.0	0.01	0.02	0.01	5.0	27.8	18.7
	(P_{bilin})	100.0	0.0	0	0.0	26.1	87.5	52.6
	(P^3_{pwl})	100.0	0.03	0.12	0.05	8.1	27.5	18.8
	(P^2_{pwl})	100.0	0.01	0.03	0.02	8.0	27.4	18.5
	(P^1_{pwl})	100.0	0.01	0.01	0.01	12.2	29.2	20.2

Table 20: Proportion of configurations solved with their VBS, CT and AE statistics for the 1-HUC for different number of units (instances A-N-1 to A-N-3 and B-N-1 to B-N-3)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
N=2	$(P_{5PL-max})$	100.0	103.34	186.38	144.86	0.2	0.4	0.3
	$(P_{5PL-bin})$	100.0	15.28	28.47	21.88	0.2	0.4	0.3
	$(P_{2D-poly})$	100.0	0.43	0.69	0.56	0.8	3.9	2.4
	(P_{op})	100.0	1.47	2.25	1.86	0.3	0.4	0.3
	(P_{bilin})	100.0	0.03	0.05	0.04	24.3	26.1	25.2
	(P^3_{pwl})	100.0	0.4	7.06	3.73	1.3	3.5	2.4
	(P^2_{pwl})	100.0	0.09	0.18	0.14	5.8	6.5	6.2
	(P^1_{pwl})	100.0	0.01	0.02	0.01	11.8	67.9	39.9
N=4	$(P_{5PL-max})$	100.0	1764.35	5988.44	3876.39	0.1	0.4	0.2
	$(P_{5PL-bin})$	100.0	13.75	63.73	38.74	0.1	0.4	0.2
	$(P_{2D-poly})$	100.0	2.76	3.47	3.12	0.4	3.1	1.8
	(P_{op})	100.0	0.49	1.57	1.03	0.0	0.3	0.1
	(P_{bilin})	100.0	0.04	0.05	0.04	19.6	22.1	20.9
	(P^3_{pwl})	100.0	0.75	5.73	3.24	0.3	3.3	1.8
	(P^2_{pwl})	100.0	0.11	0.15	0.13	4.0	6.4	5.2
	(P^1_{pwl})	100.0	0.02	0.02	0.02	12.0	70.9	41.5
N=6	$(P_{5PL-max})$	100.0	1055.49	1055.49	1055.49	0.0	0	0.0
	$(P_{5PL-bin})$	100.0	7.63	36.01	21.82	0.0	0.3	0.1
	$(P_{2D-poly})$	100.0	1.11	3.81	2.46	0.7	8.7	4.7
	(P_{op})	100.0	0.38	1.3	0.84	0.0	0.9	0.5
	(P_{bilin})	100.0	0.04	0.04	0.04	13.6	20.4	17.0
	(P^3_{pwl})	100.0	0.5	6.63	3.56	0.4	4.2	2.3
	(P^2_{pwl})	100.0	0.14	0.22	0.18	3.7	6.4	5.1
	(P^1_{pwl})	100.0	0.02	0.02	0.02	10.6	75.8	43.2

Table 21: Proportion of configurations solved with their VBS, CT and AE statistics for the fixed-head 1-HUC for different number of units (instances A-N-1 to A-N-3 and B-N-1 to B-N-3)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
N=2	$(P_{5PL-max})$	100.0	0.11	0.12	0.11	10.7	20.6	15.7
	$(P_{5PL-bin})$	100.0	0.25	0.4	0.33	10.7	20.6	15.7
	$(P_{2D-poly})$	100.0	0.06	0.08	0.07	14.6	27.3	20.9
	(P_{op})	100.0	0.0	0.01	0.01	10.3	21.0	15.7
	(P_{bilin})	100.0	0.0	0	0.0	41.0	66.6	53.8
	(P^3_{pwl})	100.0	0.02	0.02	0.02	20.2	21.9	21.0
	(P^2_{pwl})	100.0	0.01	0.01	0.01	19.9	21.8	20.9
	(P^1_{pwl})	100.0	0.01	0.01	0.01	21.9	22.2	22.0
N=4	$(P_{5PL-max})$	100.0	0.11	0.33	0.22	12.2	24.9	18.5
	$(P_{5PL-bin})$	100.0	0.14	0.76	0.45	14.4	22.4	18.4
	$(P_{2D-poly})$	100.0	0.06	0.09	0.07	10.3	11.2	10.8
	(P_{op})	100.0	0.01	0.01	0.01	11.6	22.0	16.8
	(P_{bilin})	100.0	0.0	0	0.0	36.7	56.7	46.7
	(P^3_{pwl})	100.0	0.02	0.03	0.03	22.7	23.4	23.0
	(P^2_{pwl})	100.0	0.01	0.01	0.01	22.8	23.0	22.9
	(P^1_{pwl})	100.0	0.01	0.01	0.01	21.7	26.5	24.1
N=6	$(P_{5PL-max})$	100.0	0.1	0.33	0.22	13.0	25.4	19.2
	$(P_{5PL-bin})$	100.0	0.03	0.82	0.42	24.7	90.2	57.5
	$(P_{2D-poly})$	100.0	0.06	0.07	0.07	13.9	15.6	14.8
	(P_{op})	100.0	0.0	0	0.0	10.3	24.0	17.1
	(P_{bilin})	100.0	0.0	0	0.0	35.0	48.4	41.7
	(P^3_{pwl})	100.0	0.02	0.03	0.03	23.1	25.1	24.1
	(P^2_{pwl})	100.0	0.01	0.01	0.01	23.5	24.9	24.2
	(P^1_{pwl})	100.0	0.01	0.01	0.01	22.2	23.6	22.9

Table 22: Proportion of configurations solved with their VBS, CT and AE statistics for the 1-HUC for small and large volumes (instances A-S-1, A-S-2, B-S-1 and B-S-2)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
$S \in \{0.2, 0.36\}$	$(P_{5PL-max})$	100.0	103.34	186.38	144.86	0.2	0.4	0.3
	$(P_{5PL-bin})$	100.0	15.28	28.47	21.88	0.2	0.4	0.3
	$(P_{2D-poly})$	100.0	0.43	0.69	0.56	0.8	3.9	2.4
	(P_{op})	100.0	1.47	2.25	1.86	0.3	0.4	0.3
	(P_{bilin})	100.0	0.03	0.05	0.04	24.3	26.1	25.2
	(P^3_{pwl})	100.0	0.4	7.06	3.73	1.3	3.5	2.4
	(P^2_{pwl})	100.0	0.09	0.18	0.14	5.8	6.5	6.2
	(P^1_{pwl})	100.0	0.01	0.02	0.01	11.8	67.9	39.9
$S \in \{0.002, 0.0036\}$	$(P_{5PL-max})$	100.0	123.58	131.2	127.39	0.5	0.8	0.7
	$(P_{5PL-bin})$	100.0	7.2	16.72	11.96	0.5	0.8	0.7
	$(P_{2D-poly})$	100.0	0.26	0.36	0.31	3.7	17.5	10.6
	(P_{op})	100.0	0.18	0.43	0.3	0.5	1.0	0.8
	(P_{bilin})	100.0	0.05	0.06	0.06	9.6	29.8	19.7
	(P^3_{pwl})	100.0	0.05	0.11	0.08	2.7	171.9	87.3
	(P^2_{pwl})	100.0	0.01	0.01	0.01	10.8	173.6	92.2
	(P^1_{pwl})	100.0	0.01	0.02	0.01	86.3	100.9	93.6

Table 23: Proportion of configurations solved with their VBS, CT and AE statistics for the fixed-head 1-HUC for small and large volumes (instances A-S-1, A-S-2, B-S-1 and B-S-2)

Instances	Model	%S	min-CT	max-CT	avg-CT	min-AE	max-AE	avg-AE
$S \in \{0.2, 0.36\}$	$(P_{5PL-max})$	100.0	0.11	0.12	0.11	10.7	20.6	15.7
	$(P_{5PL-bin})$	100.0	0.25	0.4	0.33	10.7	20.6	15.7
	$(P_{2D-poly})$	100.0	0.06	0.08	0.07	14.6	27.3	20.9
	(P_{op})	100.0	0.0	0.01	0.01	10.3	21.0	15.7
	(P_{bilin})	100.0	0.0	0	0.0	41.0	66.6	53.8
	(P^3_{pwl})	100.0	0.02	0.02	0.02	20.2	21.9	21.0
	(P^2_{pwl})	100.0	0.01	0.01	0.01	19.9	21.8	20.9
	(P^1_{pwl})	100.0	0.01	0.01	0.01	21.9	22.2	22.0
$S \in \{0.002, 0.0036\}$	$(P_{5PL-max})$	100.0	0.14	0.18	0.16	0.1	1.2	0.7
	$(P_{5PL-bin})$	100.0	0.72	0.84	0.78	0.2	1.0	0.6
	$(P_{2D-poly})$	100.0	0.09	0.09	0.09	4.2	16.7	10.4
	(P_{op})	100.0	0.0	0	0.0	0.2	1.1	0.7
	(P_{bilin})	100.0	0.0	0	0.0	9.8	30.3	20.1
	(P^3_{pwl})	100.0	0.03	0.03	0.03	0.8	176.3	88.6
	(P^2_{pwl})	100.0	0.01	0.01	0.01	0.8	175.8	88.3
	(P^1_{pwl})	100.0	0.01	0.01	0.01	0.7	177.8	89.2