Graphical Abstract

Neural network informed day-ahead scheduling of pumped hydro energy storage

Pietro FAVARO, Mihály DOLÁNYI, François VALLÉE, Jean-François TOUBEAU

Highlights

Neural network informed day-ahead scheduling of pumped hydro energy storage

Pietro FAVARO, Mihály DOLÁNYI, François VALLÉE, Jean-François TOUBEAU

- The modelling accuracy of efficiency curves affects the market value of pumped hydro
- Neural networks accurately model the unit's efficiency curves
- Neural-network-informed optimization can be cast as a mixed-integer linear problem
- Pruning neural network weights decreases the computation time of the optimization
- Neural-network-informed optimization ensures the safe provision of flexibility

Neural network informed day-ahead scheduling of pumped hydro energy storage

Pietro FAVARO[a,1,*], Mihály DOLÁNYI[a], François VALLÉE[a], Jean-François TOUBEAU[b]

[a] *Power System and Market Research Group (PSMR), University of Mons, Boulevard Dolez 31, Mons, 7000, Wallonia, Belgium*
[b] *Energy Systems Integration and Modeling Group (ESIM), KU Leuven, Celestijnenlaan 300, Heverlee, 3001, Flanders, Belgium*

Abstract

This paper presents a neural network-constrained optimization model for the optimal scheduling of pumped hydro energy storage. Neural networks are trained offline to capture the complex head-dependent performance curves in both pump and turbine modes using actual operation data. The trained models are then embedded into the optimization framework that yields the optimal and physics-compliant day-ahead scheduling in energy and reserve markets for the pumped hydro energy storage. To identify the trade-off between modeling accuracy and computation burden, different neural network architectures are investigated, along with the impact of neural network sparsity, i.e., weights pruning to reduce dimensionality. The proposed approach is then compared with state-of-the-art solutions, such as piecewise linear approximations. To that end, a detailed simulator of the pumped hydro energy storage, mimicking its minute-wise behavior, is developed to accurately assess the feasibility and economic performance of the resulting schedules. Results demonstrate the ability of neural networks to better guide the optimization model, thus leading to higher profits while keeping acceptable solving times, especially when weight pruning is leveraged. In particular, we show that accurately capturing the non-linear characteristics of pumped hydro energy storage is critical to offer reliable reserve commitments to power systems.

*Corresponding author
Email address: `pietro.favaro@umons.ac.be` (Pietro FAVARO)
[1]Pietro Favaro is funded by the F.N.R.S-FRS.

Keywords: Machine learning, Neural network informed optimization, Pumped hydro energy storage, Data-driven optimization, Day-ahead scheduling

---

Nomenclature

The various notations used in the paper are described in this section.

A. Sets and Indices

T	Set of time steps, index t.
H	Set of PHES plants, index h.
R	Set of the reserve products, index r.
$R^{+} \subseteq R$	Set of the upward reserve products.
$R^{-} \subseteq R$	Set of the downward reserve products.
f_u, f_d	Frequency containment reserve upward and downward.
a_u, a_d	Automatic frequency restoration reserve upward and downward.
T, P	Turbine and pump modes, index i.

B. Variables

res_r	Total reserve capacity committed to the reserve product r, [MW].
$res^{i}_{h,r}$	Reserve capacity committed by each plant h to a given reserve product r in mode i, [MW].
$v^{\mathrm{res}}_{h,t,r}$	Additional water volume moved because of the reserve activation of plant h at time step t for the reserve product r, [m^3].
e^{DA}_{t}	Energy position in the day-ahead energy-only market at time step t, [MWh].
$z^{\mathrm{T}}_{h,t}, z^{\mathrm{P}}_{h,t}$	Binary variable indicating the operating mode, i.e., turbine (T) or pump (P) of plant h at time step t.

$p^{\mathrm{T}}_{h,t}, p^{\mathrm{P}}_{h,t}$	Output power in turbine (T) and pump (P) modes of plant h at time step t, [MW].
$\underline{p}^{i}_{h,t}, \overline{p}^{i}_{h,t}$	Minimum and maximum output power of plant h at time step t in mode i, [MW].
$q^{\mathrm{T}}_{h,t}, q^{\mathrm{P}}_{h,t}$	Water flow rate in turbine (T) and pump (P) modes of plant h at time step t, [m^3/s].
$v^{\mathrm{up}}_{h,t}, v^{\mathrm{low}}_{h,t}$	Water volume in the upper (up) and lower (low) reservoirs of plant h at time step t, [m^3].
$h^{\mathrm{up}}_{h,t}, h^{\mathrm{low}}_{h,t}$	Water head in the upper (up) and lower (low) reservoirs of plant h at time step t, [m].
$h^{\mathrm{net}}_{h,t}$	Net water head of plant h at time step t, [m].
$h^{\mathrm{loss}}_{h,t}$	Water head loss of plant h at time step t, [m].
$b^{(l)}_{k}$	Binary variable associated with the activation function of the neuron k in layer l.
$\hat{y}^{(l)}$	Input vector of layer (l) of a given NN.
$\hat{y}^{(l)}_{k}$	One element of the vector $\hat{y}^{(l)}$.
$y^{(l)}_{k}$	Output vector of layer (l) of a given NN.
$y^{(l)}_{k}$	One element of the vector $\hat{y}^{(l)}$.
$A^{i}_{h,t}, B^{i}_{h,t}$	Additional variables to enforce the consistency of the NN constraints associated with the UPC of mode i for plant h at time step t.

C. Parameters

Δt	Time step duration of the optimisation [h].
$\lambda^{\mathrm{res}}_{r}$	Price for the reserve capacity made available in each product r, [€/MW].
$\lambda^{\mathrm{DA}}_{t}$	Electricity price on the energy-only day-ahead market at time step t, [€/MWh].
C^{op}_{h}	Operating cost of plant h, [€/MWh].
$\Delta R^{i}_{h,r}$	Ramping ability of plant h in mode i for reserve product r, [MW].
$\overline{Q}^{i}_{h}$	Maximum water flow rate in mode i for plant h, [m^3/s].

$\overline{V}_h^{\mathrm{up}}, \overline{V}_h^{\mathrm{low}}$	Maximum water volume of the upper (up) and lower (low) reservoirs of the plant h, [m^3].
V_h^{target}	Target volume of water in the upper reservoir at the end of the optimization horizon for plant h, [m^3].
$\underline{H}_h^{\mathrm{net}}, \overline{H}_h^{\mathrm{net}}$	Minimum and maximum net water head of plant h.
$W^{(l)}$	Weight matrix associated with layer l.
$B^{(l)}$	Bias vector associated with layer l.
$\underline{M1}_h^i, \overline{M1}_h^i, \underline{M2}_h^i, \overline{M2}_h^i$	Big-M constants of plant h in mode i.

D. Functions

$f_{h,t}^{\mathrm{up}}, f_{h,t}^{\mathrm{low}}$	Geometry of the upper (up) and lower (low) reservoirs of plant h at time step t.
$f_h^{\mathrm{UPC},i}$	Unit Performance Curve (UPC) of the hydraulic machine of plant h operating in mode i.
$\underline{p}_h^i, \overline{p}_h^i$	Upper and lower power bounds of the UPC associated with mode i of plant h.

1. Introduction

In order to mitigate uncertainties from variable renewable generation, power systems need to rely on increasing amounts of flexibility, which can be provided by energy storage. For instance, the European Union assessed a need for 97 GW of additional storage capacity by 2030 in order to support its energy policies [1]. Pumped-Hydro Energy Storage (PHES) is a mature and robust technology which currently represents more than 95% of the worldwide utility-scale storage capacity [2]. In recent years, PHES has known major technological improvements, driven by enhancements in power electronics that allow the units to operate at variable speeds [3]. Micro-PHES can also benefit from these advancements, as demonstrated in [4]. The resulting increased operating range, and associated flexibility, enables a higher penetration of renewable-based fluctuating generation [5]. In that direction, different approaches have been developed. First, variable-speed technologies enable to provide balancing services to the power system [6]. Second, flexibility can be created by coupling constant-speed PHES with energy storage technologies such as batteries or flywheels [7].

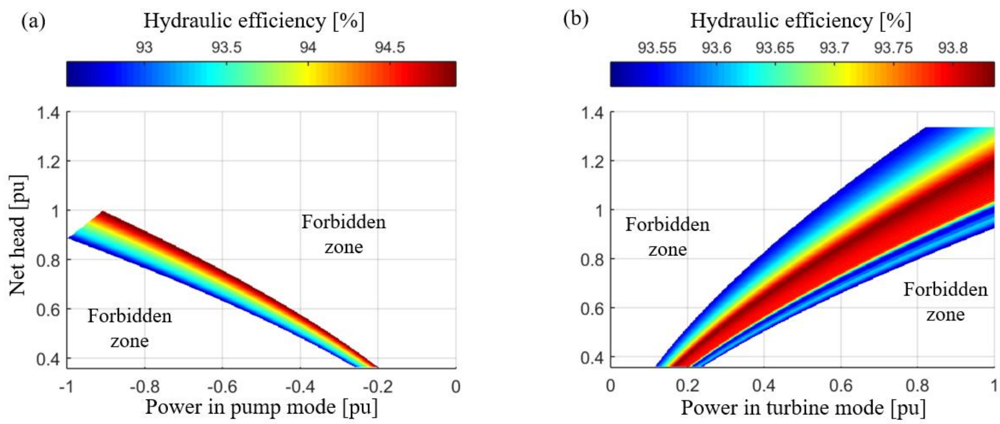

Figure 1: Unit Performance Curve (UPC) of a conventional Francis machine in (a) pump mode, (b) turbine mode.

In this paper, the multi-period day-ahead scheduling problem faced by PHES plant owners, who participate in both energy and reserve markets, is studied. In addition to conventional constraints (e.g., ramping limits, power and energy bounds), the constraints associated with the operation of hydraulic units must also be considered. In particular, the operating characteristics of the hydraulic machine, referred to as Unit Performance Curves (UPC), need to be properly described. The UPC is a three-dimensional relationship between the net head (i.e., the height difference between the water levels in the upper and lower basins, minus the losses), the output power and the water flow rate. Figure 1 depicts the typical UPCs for a variable-speed Francis pump-turbine, i.e., a reversible technology that can either be used as a pump or a turbine over a wide range of head-level. The UPCs were obtained through laboratory measurements of a reduced-scale model [8]. This work was part of the SmartWater project [9]. The curves were obtained via polynomial interpolation of the measurements aiming at minimizing the squared modeling error.

Figure 1 shows that the UPCs are defined only in some parts of the net head/power plane, rendering the underlying optimization problem discontinuous. Indeed, when the water flow rate is too high, cavitation (i.e., rapid implosion of gaseous cavities) can appear and damage the hydraulic machine. In contrast, a low water flow may trigger harmful mechanical vibrations through self-excitation and severe system erosion [10]. Therefore, accurately representing the head-dependent UPCs is key to guarantee the feasibility, reliability and cost-efficiency of the PHES operating schedule.

Various optimization models with different objective functions are described in the literature for short-term scheduling of hydro generation. In [11], the efficiency of hydro-power plant is maximized, thereby maximizing the energy production on the long run. Other approaches aim at maximizing the profits. With the aim of guaranteeing the highest profits possible in case of extreme scenarios of uncertainties, robust optimization is employed in reference [12]. In [13], the day-ahead scheduling of PHES is performed under considerations for irrigation systems and uncertain wind power. In those formulations, the PHES dynamics are modelled using deterministic equations. However, in some installations such as underground PHES, stochastic modelling is more appropriate because of the water exchanges with nearby aquifers and rock porosity [14]. All resulting models are challenging to solve. A straightforward approach is to rely on non-linear optimization, but the solvers typically suffer from a high computational burden and offer no optimality guarantee on the solution [15]. This problem is further exacerbated when considering uncertainties due to the increased dimensionality of the resulting model [16]. To improve computational efficiency, alternative strategies relying on Lagrangian relaxation were also tested but they exhibit convergence issues [17]. Meta-heuristic algorithms, another family of optimization algorithms which are agnostic to the problem structure, were then employed. Genetic algorithm, which mimics the genetic evolution of a population through breeding and mutation to try converging towards the optimal solution, was used for cost minimization [18]. It was also applied to a more complex multi-objective optimization scheme including environmental impact [19]. Particle Swarm Optimization (PSO), which is based on social behaviour of animal populations such as fish schooling or bird flocking, was used for optimizing the PHES operation with one or several upper basins [20]. However, such methods do not offer any information on the quality of the final solution which is dependent on the starting point of the algorithm due to the inherent stochastic nature of the process. Dynamic programming has also been applied but it was proven to be very sensitive to the curse of dimensionality [21]. This remains valid even when approximating the hydropower generation function with a concave curve [22].
Alternatively, piecewise linear approximation of UPCs, which relies on Mixed-Integer Linear Programming (MILP), has been widely investigated [23]. The discontinuity between the pump and the turbine UPCs can be modelled using binaries [24]. Each UPC can be approximated with nine triangle plane

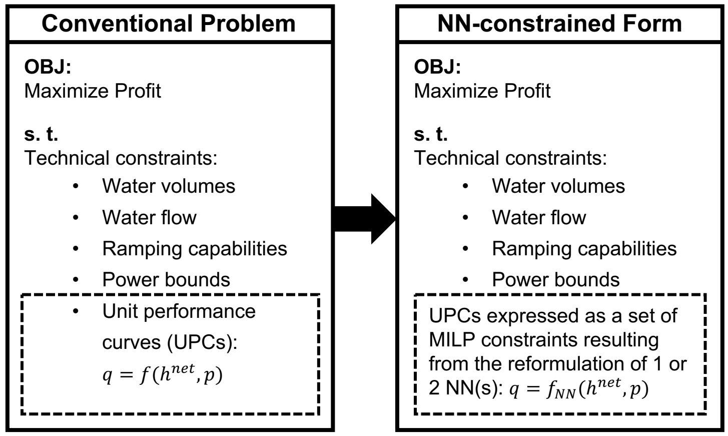

Figure 2: Proposed approach with NN-constrained optimization

pieces [25]. The piecewise linear approximation was extended to 25 triangle pieces in [26]. In [27], the authors modelled the non-linearity of the generation function in the case of hydro-power. A chance-constraint optimization for underground PHES on top of a MILP approximation of the UPCs based on rectangle plane pieces is proposed in [14]. Those formulations have been improving continuously, becoming more compact and tighter. Nevertheless, the number of pieces used to model the non-linear curves remains an expert-based decision, which results from a trade-off between the modeling accuracy and the number of additional variables. Furthermore, this segmentation of the input space is usually performed in a uniform fashion. Hence, it may arise that a linear part of a curve benefits from the same modeling granularity as a highly non-linear area, resulting in a poor allocation of computing resources. Finally, the solution of the MILP algorithm may not always be feasible due to these inherent approximations [28].

To address the limitations of conventional piecewise-linear methods, this paper introduces Neural Networks (NNs) to represent the nonlinear UPCs. Instead of partitioning the feasible space arbitrarily, the proposed method leverages the modeling power of NNs, defined with piecewise linear activation functions, e.g., Rectified Linear Unit (ReLU), in order to obtain an optimal approximation of the UPCs for a minimal number of neurons [2]. The *exact*

[2]The resulting NN architecture is a trade-off between the precision of the UPCs approximation and the number of variables (continuous and binary) introduced in the optimization

mixed-integer reformulation of the NN can then be integrated into the PHES optimization problem which can be solved using an off-the-shelf MILP solver [29]. Figure 2 summarizes the proposed data-driven approach.

The main contributions of this paper can be summarized as follows:

(a) We present a novel data-driven NN approximation of the UPCs to model the non-linear behaviour of a pumped-hydro energy storage unit. Then, an exact mixed-integer reformulation is developed to embed the trained NN architecture into the optimization model which yields the day-ahead scheduling in energy and reserve markets.

(b) We develop a new modeling strategy to enforce the formulation consistency between the PHES operation modes (turbine, pump, idle). This is achieved by ensuring that, in a given operation mode, only the NN constraints associated with this mode are binding.

(c) We improve the compactness of the resulting MILP reformulation by investing two complementary strategies. First, we model both the pump and the turbine UPCs using a single NN. Second, we investigate sparse neural networks, i.e., by pruning some weights to decrease the number of parameters.

The effectiveness of the proposed method is compared with state-of-the-art approaches such as uniformly segmented piecewise linear approximations and linear regression, using data of a candidate site for a PHES plant located in Maizeret, Belgium, is used. The quality of the decisions is assessed through a detailed simulator of the PHES operation in energy and reserve market floors in order to obtain a fair and accurate estimation of ex-post profits. Thirty scenarios are used to assess the ex-post profit performance of the models. This validation is completed with a sensitivity analysis on the ex-post profit with respect to the uncertain market penalties.

The rest of the paper is organized as follows. The formulation of the day-ahead scheduling problem, with a joint participation in energy and reserve markets, is presented in Section 2. Section 3 describes the NN modeling of the UPCs and their reformulation into MILP constraints. Section 4 discusses and compares the results of the proposed approach with state-of-the-art bench-

---

problem when reformulating exactly the NNs as constraints.

marks. Lastly, Section 5 provides some conclusions and perspectives for future research.

2. Problem Formulation

In this section, the formulation of the day-ahead scheduling problem faced by a PHES operator, based on [14], is presented. We assume that the day-ahead energy and reserve markets are jointly cleared for the 24 hours of the next day. The formulation adopts a price-taker approach with a perfect price forecast.

2.1. Objective Function

The objective of the PHES operator is to maximize their profit on the day-ahead energy and reserve markets, subject to the operational and market constraints. To do so, the decisions to pump (i.e., consuming power) or to operate in turbine mode (i.e., generating power) are placed as quantity orders on the markets in the most profitable way. The objective function of the day-ahead scheduling problem (1), which maximizes the operator's expected profit, is made up of (i) the benefits earned by providing a constant amount of upward and/or downward reserve capacities res_r over the 24 hourly market periods of the next day for each type of reserve product r at price λ_r^{res}; (ii) the inter-temporal arbitrage revenues on the day-ahead market obtained by selling or purchasing an amount of electricity e_t^{DA} at price λ_t^{DA} when operating in turbine or pump mode; (iii) the operating cost, incurred by each unit h, is assumed to be linear and proportional to the energy produced by the turbine $p_{h,t}^{\mathrm{T}} \cdot \Delta t$ and consumed by the pump $p_{h,t}^{\mathrm{P}} \cdot \Delta t$ with a constant operating cost C_h^{op}.

$$\Phi = \sum_{r \in R} \underbrace{24 \cdot \lambda_r^{\mathrm{res}} \cdot res_r}_{(i)} + \sum_{t \in T} \left[\underbrace{\lambda_t^{\mathrm{DA}} \cdot e_t^{\mathrm{DA}}}_{(ii)} - \sum_{h \in H} \underbrace{C_h^{\mathrm{op}} \cdot (p_{h,t}^{\mathrm{T}} + p_{h,t}^{\mathrm{P}}) \cdot \Delta t}_{(iii)} \right] \quad (1)$$

Objective function (1) involves three assumptions. Firstly, the benefits from the reserve activation are supposed to offset exactly to the PHES operating costs. Second, the operator cannot willingly deviate from his day-ahead schedule, thus explaining why imbalance penalties are absent from the objective function. Lastly, the operational costs scale linearly with the electrical energy generated and consumed. This is a necessary simplification of the very complex reality. For a more realistic approximation of the underlying operational costs, the fluid dynamics in the hydraulic machine need to be modelled but there exists no clear consensus on an effective model to evaluate the resulting wear and tear costs [30].

2.2. Energy Balance

For each hourly market period, the energy offered to the market either in pump mode or turbine mode must be delivered by the PHES plant, which is ensured by constraint (2).

$$e_t^{\mathrm{DA}} = \Delta t \cdot \sum_{h \in H} \left(p_{h,t}^{\mathrm{T}} - p_{h,t}^{\mathrm{P}} \right) \qquad \forall t \tag{2}$$

2.3. Reserve Allocation Constraints

The PHES flexibility can also be valued and traded on the reserve market. The goal is to identify the profit-maximizing allocation of reserve products $r \in R$ over the daily scheduling horizon.

The Transmission System Operator (TSO) is in charge of sizing, allocating and activating the reserves. These reserves are divided into downward and upward products. The downward reserve is activated when the grid frequency is too high and can be provided either by reducing the generated power in turbine mode or by increasing the consumed power in pump mode. On the contrary, the upward reserve is activated when the grid frequency is too low and can be provided either by increasing the generated power in turbine mode or by reducing the consumption in pump mode.

Furthermore, both the downward and the upward reserves are organized in three products, which differ in their expected response speed and price. Frequency Containment Reserves (FCRs) are the first to be automatically called upon in the emergence of a contingency, and must be fully operational within 30 seconds. Then, the automated Frequency Restoration Reserves (aFRRs), which need to respond within 7.5 minutes, are activated in order to free up the capacities of the FCRs for future contingencies. Lastly, if the imbalance persists, the TSO requests the dispatch of manual Frequency Restoration Reserves (mFRRs) that remain online until the resolution of the disturbance.

At each time step, the committed capacity of the PHES plant in the six reserve products is the sum of the flexibility available from all units (3).

$$\sum_{h\in H}\left(res^{\mathrm{T}}_{h,t,r}+res^{\mathrm{P}}_{h,t,r}\right)=res_r \qquad \forall t,r \tag{3}$$

The PHES participation to the reserve market is limited by its ramping ability. Hence, it must be ensured that the PHES ramping ability is not doubly allocated to the different reserve products [14], which is enforced by constraints (4) - (8).

$$res^i_{h,t,r}\le z^i_{h,t}\cdot\Delta R^i_{h,r} \qquad \forall\ h,t,i\in\{\mathrm{T},\mathrm{P}\},r\in\{f_u,f_d\} \tag{4}$$

$$res^i_{h,t,f_d}+res^i_{h,t,a_d}\le z^i_{h,t}\cdot\Delta R^i_{h,a_d} \qquad \forall h,t,i\in\{\mathrm{T},\mathrm{P}\} \tag{5}$$

$$res^i_{h,t,f_u}+res^i_{h,t,a_u}\le z^i_{h,t}\cdot\Delta R^i_{h,a_u} \qquad \forall h,t,i\in\{\mathrm{T},\mathrm{P}\} \tag{6}$$

$$\sum_{r\in R^-}res^i_{h,t,r}\le z^i_{h,t}\cdot\Delta R^i_{h,m_d} \qquad \forall h,t,i\in\{\mathrm{T},\mathrm{P}\} \tag{7}$$

$$\sum_{r\in R^+}res^i_{h,t,r}\le z^i_{h,t}\cdot\Delta R^i_{h,m_u} \qquad \forall h,t,i\in\{\mathrm{T},\mathrm{P}\} \tag{8}$$

2.4. Technical Constraints

The state-of-charge of the PHES plant is a function of the water volume in the lower and upper basins. Those volumes cannot be lower than a given threshold ($\underline{V}^{\mathrm{low}}_h$ and $\underline{V}^{\mathrm{up}}_h$), nor can they be higher than the maximum content of the reservoirs ($\overline{V}^{\mathrm{low}}_h$ and $\overline{V}^{\mathrm{up}}_h$). Moreover, the PHES scheduling involving the provision of upward reserve must ensure that the upper reservoir contains enough water and the lower reservoir is able to receive the potential surplus of water inflow (9). Reversely for downward reserve, there must be a sufficient volume of water in the lower basin, along with enough free space in the upper one to accommodate the extra water transferred (10).

$$\underline{V}^{\mathrm{up}}_h+\sum_{t'=1}^{t}\sum_{r\in R^-}v^{\mathrm{res}}_{h,t',r}\le v^{\mathrm{low}}_{h,t}\le\overline{V}^{\mathrm{up}}_h-\sum_{t'=1}^{t}\sum_{r\in R^-}v^{\mathrm{res}}_{h,t',r} \qquad \forall h,t \tag{9}$$

$$\underline{V}^{\mathrm{low}}_h+\sum_{t'=1}^{t}\sum_{r\in R^+}v^{\mathrm{res}}_{h,t',r}\le v^{\mathrm{low}}_{h,t}\le\overline{V}^{\mathrm{low}}_h-\sum_{t'=1}^{t}\sum_{r\in R^+}v^{\mathrm{res}}_{h,t',r} \qquad \forall h,t \tag{10}$$

The amount of water linked to the activation of reserve is determined as follows

$$v^{\text{res}}_{h,t',r} = \frac{3600 \cdot 10^6 \cdot \left(res^{\text{T}}_{h,t',r} + res^{\text{P}}_{h,t',r}\right)}{\eta_{h,t'} \cdot \rho \cdot g \cdot h^{\text{net}}_{h,t'}} \tag{11}$$

where $g = 9.81$ m/s^2, ρ is the volumetric mass density of water (1000 kg/m^3). In (11), nonlinearities arise from the efficiency $\eta_{h,t'}$ and the net head available $h^{\text{net}}_{h,t'}$. Here, conservative values are chosen, thus overestimating the volume of water necessary for reserves $v^{\text{res}}_{h,t',r}$ and hedging against infeasible states.

The water volume within each reservoir at a time step t is obtained based on the water volumes at the previous time step $t-1$ and the current water discharge $q^i_{h,t}$, as shown by (12) and (13).

$$v^{\text{up}}_{h,t} = v^{\text{up}}_{h,t-1} + \left(q^{\text{P}}_{h,t} - q^{\text{T}}_{h,t}\right) \cdot \Delta t \qquad \forall h, t \tag{12}$$

$$v^{\text{low}}_{h,t} = v^{\text{low}}_{h,t-1} + \left(q^{\text{T}}_{h,t} - q^{\text{P}}_{h,t}\right) \cdot \Delta t \qquad \forall h, t \tag{13}$$

Each PHES unit has three distinct operating states, i.e., idle, pump and turbine modes. Only one mode is accessible over a given time period t, which is enforced by (14). Then, the water discharge is upper bounded by Equations (15) - (16).

$$z^{\text{P}}_{h,t} + z^{\text{T}}_{h,t} \leq 1 \qquad \forall h, t, z^{\text{i}}_{h,t} \in \{0, 1\} \tag{14}$$

$$q^{\text{P}}_{h,t} \leq z^{\text{P}}_{h,t} \cdot \overline{q_h}^{\text{P}} \qquad \forall h, t \tag{15}$$

$$q^{\text{T}}_{h,t} \leq z^{\text{T}}_{h,t} \cdot \overline{q_h}^{\text{T}} \qquad \forall h, t \tag{16}$$

At the end of the scheduling horizon, the water stored in the upper reservoir represents future financial values. The target amount of water to retain in the upper reservoir at the end of the horizon $t = T$ is set as a threshold using (17).

$$v^{\text{up}}_{h,t=T} \geq V^{\text{target}}_h \qquad \forall h \tag{17}$$

The net head depends on the water level in both reservoirs, which are themselves dependent on the water volumes through the reservoir geometry (18) and (19). Due to the friction of the fluid against the penstock walls and the inner turbulence losses, the net head is lower than the gross head (21). This head loss is usually modeled as proportional to the square of the water discharge (20) [31].

$$h_{h,t}^{\mathrm{up}} = f_{h,t}^{\mathrm{up}}\left(v_{h,t}^{\mathrm{up}}\right) \qquad \forall h, t \tag{18}$$

$$h_{h,t}^{\mathrm{low}} = f_{h,t}^{\mathrm{low}}\left(v_{h,t}^{\mathrm{low}}\right) \qquad \forall h, t \tag{19}$$

$$h_{h,t}^{\mathrm{loss}} = c_{h}^{\mathrm{loss}} \cdot \left(q_{h,t}^{\mathrm{T}} + q_{h,t}^{\mathrm{P}}\right)^{2} \qquad \forall h, t \tag{20}$$

$$h_{h,t}^{\mathrm{net}} = \underbrace{h_{h,t}^{\mathrm{up}} - h_{h,t}^{\mathrm{low}}}_{\text{gross head}} - h_{h,t}^{\mathrm{loss}} \qquad \forall h, t \tag{21}$$

Equations (22) and (23) enforce the power bounds of the feasible operating zones while considering the possible reserve participation.

$$z_{h,t}^{\mathrm{P}} \cdot \underline{p}_{h,t}^{\mathrm{P}} + \sum_{r\in R^{+}} res_{h,r}^{\mathrm{P}} \leq p_{h,t}^{\mathrm{P}} \leq z_{h,t}^{\mathrm{P}} \cdot \overline{p}_{h,t}^{\mathrm{P}} - \sum_{r\in R^{-}} res_{h,r}^{\mathrm{P}} \qquad \forall h, t \tag{22}$$

$$z_{h,t}^{\mathrm{T}} \cdot \underline{p}_{h,t}^{\mathrm{T}} + \sum_{r\in R^{-}} res_{h,r}^{\mathrm{T}} \leq p_{h,t}^{\mathrm{T}} \leq z_{h,t}^{\mathrm{T}} \cdot \overline{p}_{h,t}^{\mathrm{T}} - \sum_{r\in R^{+}} res_{h,r}^{\mathrm{T}} \qquad \forall h, t \tag{23}$$

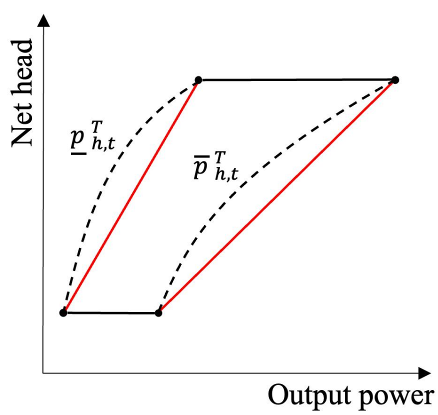

Figure 3: Domain modeling of the UPC in turbine mode using a trapezoid.

Finally, the turbine/pump operating point at each time step must comply with the unit performance curve (UPC). The UPC is a three-dimensional non-convex non-concave relationship between the net head, the ouptut power and the water discharge. Both pump and turbine modes are characterized by their own UPC curvature (24) and UPC domain (Figure 1). The UPC domains must be respected since the hydraulic machine risks to be unstable out of these zones due to damaging mechanical vibrations or cavitation effects. The safe power bounds of the UPC domain $\underline{p}^i_{h,t}$ and $\overline{p}^i_{h,t}$ vary non-linearly with respect to the net head $h^{\text{net}}_{h,t}$. To keep the problem tractable and avoid a non-linear formulation, each domain is approximated by a trapezoid, as shown in Figure 3, following Constraints (25) and (26).

$$p^i_{h,t} = f^{\text{UPC,i}}_h \left(q^i_{h,t}, h^{\text{net}}_{h,t}\right) \qquad \forall h, t, i \in \{\text{T}, \text{P}\} \tag{24}$$

$$\underline{p}^i_{h,t} = \frac{\underline{p}^i_h(\overline{H}^{\text{net}}_h) - \underline{p}^i_h(\underline{H}^{\text{net}}_h)}{\overline{H}^{\text{net}}_h - \underline{H}^{\text{net}}_h} \cdot h^{\text{net}}_{h,t} + \underline{p}^i_h(\underline{H}^{\text{net}}_h) \qquad \forall h, t, i \in \{\text{T}, \text{P}\} \tag{25}$$

$$\overline{p}^i_{h,t} = \frac{\overline{p}^i_h(\overline{H}^{\text{net}}_h) - \overline{p}^i_h(\underline{H}^{\text{net}}_h)}{\overline{H}^{\text{net}}_h - \underline{H}^{\text{net}}_h} \cdot h^{\text{net}}_{h,t} + \overline{p}^i_h(\underline{H}^{\text{net}}_h) \qquad \forall h, t, i \in \{\text{T}, \text{P}\} \tag{26}$$

3. Data-driven Reformulation of Unit Performance Curves

Traditionally, the UPCs (24) are approximated using a piecewise linear reformulation. To reduce the resulting modeling inaccuracies, we introduce a data-driven strategy relying on Neural Networks (NNs), which can be exactly reformulated as a set of mixed-integer linear constraints.

3.1. Neural Network Modeling

Neural Networks (NNs) have long demonstrated their ability to model [32] non-linear dynamic system with great accuracy. They can also be used to go further and control such complex systems [33]. Figure 4 depicts an example of feed-forward fully-connected NN. In this type of NN, the propagation of the information between two layers is described by Equation (27) where $\hat{y}^{(l)}$ is the input vector of layer l composed of $K^{(l)}$; $W^{(l)}$ is the weight matrix of size $K^{(l)} \times K^{(l-1)}$; $y^{(l-1)}$ is the output vector of size $K^{(l-1)}$ of layer $l-1$; $B^{(l)}$ is the biais vector of size $K^{(l)}$ of layer l. We can enforce sparsity on the weight matrix $W^{(l)}$ by destroying some connections between neurons (i.e., pruning weights). In this work, a weight pruning rate of 25% is imposed on certain architectures referred to as "sparse".

$$\hat{y}^{(l)} = W^{(l)} \cdot y^{(l-1)} + B^{(l)} \tag{27}$$

Each neuron takes a scalar input (an element of $\hat{y}^{(l)}$) and applies an activation function to obtain the scalar output. We selected the Rectified Linear Unit (ReLU) as the activation function since it enables to reformulate the NN as

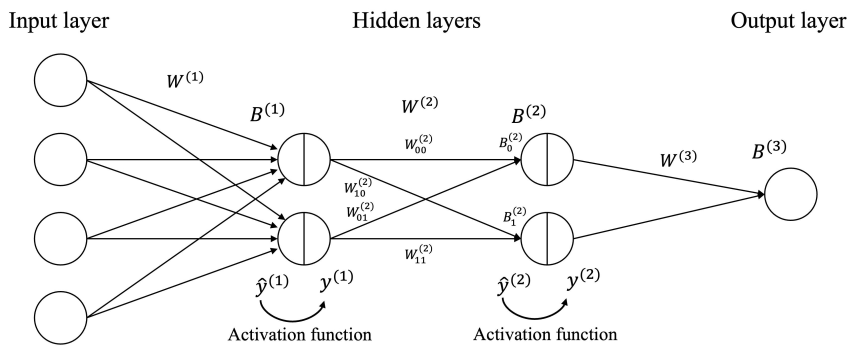

Figure 4: Feed-forward fully-connected NN with four input elements, two hidden layers with two neurons each and an output layer with one neuron.

a set of mixed-integer linear constraints. In addition, the ReLU quickens training and reaches higher accuracy for deep conventional NNs [34]. Similar observations were performed for sparse NNs [35]. The ReLU is depicted by Figure 5 and the mathematical formulation is given by (28).

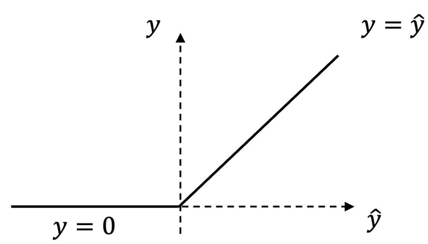

Figure 5: Rectified Linear Unit (ReLU) function.

$$y^{(l)} = \max(\hat{y}^{(l)}, 0) \tag{28}$$

The reformulation of the ReLU function requires the introduction of one binary variable b_k, two continuous scalar variables $\hat{y}_k$ and y_k, and four constraints per neuron (29)-(33) [36]. Therefore, the complexity scales with the number of neurons.

$$b_k \in \{0, 1\} \tag{29}$$

$$y_k \leq \hat{y}_k - \hat{Y}_k^{\min} \cdot (1 - b_k) \tag{30}$$

$$y_k \geq \hat{y}_k \tag{31}$$

$$y_k \leq \hat{Y}_k^{\max} \cdot b_k \tag{32}$$

$$y_k \geq 0 \tag{33}$$

The parameters $\hat{Y}_k^{\min}$ and $\hat{Y}_k^{\max}$ must be carefully chosen since they impact the tightness of the reformulation. An effective way to determine the value of these parameters is to record the minimum and maximum values of $\hat{Y}_k$ encountered during the NN training. This results in an exact, computationally efficient data-informed approximation of the underlying ReLU function.

3.2. Neural Network Formulation for Unit Performance Curve Modeling

In this subsection, the use of NNs as regressors to model the UPCs associated with the pump and turbine modes of PHES units is described. Firstly, the modeling of a single UPC is explained. Secondly, this first model is extended to account for the three operating modes (pump, turbine, idle).

A feed-forward fully-connected NN can be build to learn the UPC, using the net head and power as inputs to predict the water flow rate. Once trained, the NN is translated into constraints following Equations (27) and (29)-(33).

When the UPC of each mode is modelled with its own NN, the problem becomes infeasible. Indeed, for any pair (h^{net}, p), each NN imposes its own water discharge q, such that the model becomes inconsistent. In order to address this problem, the constraints of each NN should only be binding when the PHES is operating in the corresponding mode. In particular, the NN-constrained pump UPC should not be binding when the PHES is in turbine mode, and vice-versa.

It is thus necessary to decouple the output of the NN and the water discharge variable $q_{h,t}^i$. It can be achieved using a big-M relaxation on the NN output introducing variable $B_{h,t}^i$ (35). However, because the range of NN inputs is limited by (30) and (32), these inputs must also stay within their training range. Consequently, a big-M relaxation must also be applied on the NN inputs. Since the net head range is identical for both pump and turbine modes, only the power input must be relaxed via variable $A^i_{h,t}$ (34). The constants involved in the big-M relaxations must be chosen as small as possible to maintain the tightness of the resulting model.

$$p^i_{h,t} + \underline{M1}^i_h \cdot (1 - z^i_{h,t}) \leq A^i_{h,t} \leq p^i_{h,t} + \overline{M1}^i_h \cdot (1 - z^i_{h,t}) \qquad \forall h, t, i \in T, P \tag{34}$$

$$q^i_{h,t} + \underline{M2}^i_h \cdot (1 - z^i_{h,t}) \leq B^i_{h,t} \leq q^i_{h,t} + \overline{M2}^i_h \cdot (1 - z^i_{h,t}) \qquad \forall h, t, i \in T, P \tag{35}$$

3.3. A Single Neural Network For Joint Modeling of Pump And Turbine Unit Performance Curves

Here, we further simplify the approach by using a single NN to model both UPCs. By considering the power consumption of the pump as negative and the power generation of the turbine as positive, a single NN needs to be trained to fit both curves. Interestingly, this strategy bypasses the need to make the MILP constraints of UPC curves unbinding when the PHES is in operation. However, they must still be unbinding when the PHES is idle.

3.4. Market-based Simulator

The PHES operation schedule (obtained at the end of the optimization) may be physically infeasible, which arises from the approximations in the underlying model (e.g., on the UPCs and their bounds). In such cases, the PHES operator gets exposed to financial risks due to the need to adjust its market positions on the intraday or real-time imbalance schemes. In order to quantify those deviations in actual revenues, an advanced ex-post analysis is performed using a detailed simulator of the PHES revenue streams in the different (energy and reserve) markets, accounting for re-dispatching decisions.

The simulator used in this work takes the hourly decision output from the optimization problem and emulates the resulting PHES operations in the actual market environment. The simulator is characterized by a detailed representation of the PHES operation (with a fine timescale) and the different market floors (day-ahead and real-time stages for both energy and reserve products). Hence, it enables to accurately quantify the actual (ex-post) value of the PHES strategy.

It requires far less resources than an optimization process, which allows to drastically reduce the time granularity of the procedure. Whereas the

Algorithm 1 Simulator

Initialization of the water volumes and head levels:
$h_0^{\text{net,sim}}, h_0^{\text{up,sim}}, h_0^{\text{low,sim}}, v_0^{\text{up,sim}}, v_0^{\text{low,sim}} \leftarrow h_0^{\text{net}}, h_0^{\text{up}}, h_0^{\text{low}}, v_0^{\text{up}}, v_0^{\text{low}}$
for $t = 00{:}01$ to 24:00 with 1 minute time step **do**
  Retrieve real power range $[\underline{p}_{h,t}^{\text{sim}}, \overline{p}_t^{\text{sim}}]$ based on $\text{UPC}(h_{t-1}^{\text{net,sim}})$
  Tighten the bounds considering the reserve: $[\underline{p}_t^{\text{sim res}}, \overline{p}_t^{\text{sim res}}]$
  Set $p_t^{\text{sim}} \in [\underline{p}_t^{\text{sim res}}, \overline{p}_t^{\text{sim res}}]$ as close to p_t^{opti} as possible
  Get water flow rate $q_t^{\text{sim}} = \text{UPC}(h_t^{\text{net,sim}}, p_t^{\text{sim}})$
  if $q_t^{\text{sim}} > q_t^{\text{max}}$ **then**
    if $\{(p'_t, q'_t) \in \text{UPC} : q'_t \leq q_t^{\text{max}} \text{ and } p'_t \in [\underline{p}_t^{\text{sim res}}, \overline{p}_t^{\text{sim res}}]\} \neq \varnothing$ **then**
      $q_t^{\text{sim}} \leftarrow q'_t$ with q'_t the closest to q_t^{opti}
      $p_t^{\text{sim}} \leftarrow p'_t$
    else if $\{(p'_t, q'_t) \in \text{UPC} : q'_t \leq q_t^{\text{max}} \text{ and } p'_t \in [\underline{p}_t^{\text{sim}}, \overline{p}_t^{\text{sim}}]\} \neq \varnothing$ **then**
      $q_t^{\text{sim}} \leftarrow q'_t$ with q'_t the closest to q_t^{opti}
      $p_t^{\text{sim}} \leftarrow p'_t$
    else
      Set the machine in idle mode: $p_t^{\text{sim}}, q_t^{\text{sim}} \leftarrow 0, 0$
    end if
    if $\text{Mode}_t \neq \text{Mode}_{t-1}$ **then**
      Set the machine in idle mode: $p_t^{\text{sim}}, q_t^{\text{sim}} \leftarrow 0, 0$
    end if
  end if
  Update water volumes and head levels
end for
Compute penalties and report ex-post profit

optimization considers hourly time steps, the simulator has a time granularity of 1 minute. The state variables of the PHES simulator are indexed with the upper-script 'sim'.

As depicted in Algorithm 1, the simulator is initialized (at the start of the daily horizon) at the same values of heads and water volumes as the optimization. With the exact net head and the reserve participation, the actual available power range can be found. In this power range, the closest operating point to the one provided by the optimization is selected. Based on the power and net head values, the actual water flow rate from the UPC is retrieved. If water volume constraints are not met at the end of the minut-wise time step, the simulator selects the closest water flow rate which meets the constraints. Finally, water volumes and head values are updated, and the loop is reiterated until the end of the daily horizon. The procedure enables to quantify the exact value of the PHES scheduling, which is achieved using the following four assumptions.

1. The simulator ensures that the operator can deliver the downward and upward reserve at any time. Failure to meet this requirement leads to a penalty of 500 €/MW.

2. The non-respect of the water volumes necessary for reserve activation, which are expressed by (9) and (10), is penalized at a rate of 500 €/MWh. Equation (11) is used to convert water volumes into energy capacity.

3. If the PHES operator cannot comply with its commitments on the day-ahead energy market (DAM), electricity has to be purchased in the imbalance settlement. The Positive Price (PP) and the Negative Price (NP) attained from the Belgian TSO's website, Elia, assuming a dual pricing mechanism.

4. The lack of water in the upper reservoir at the end of the time horizon with respect to the threshold, set by (17), is penalized at a rate of 100 €/MWh. On the contrary, a surplus of water is valued at a rate of 60.22 €/MWh, i.e., the daily average price of the MWh on the day-ahead market, since it can be traded the following day.

4. Case Study

In this section, the two proposed approaches, i.e., using either a different NN for each operation mode or a single NN for both modes, are applied to the day-ahead scheduling problem of a fictitious PHES plant located in Maizeret, Belgium. After presenting the data in subsection 4.1, subsection 4.2 discusses the training outcome of the selected NN architectures. The NN models are then embedded within the optimization and the results are presented and analyzed in subsection 4.3. Both NN frameworks described in section 3 are compared with a conventional uniformly segmented piecewise linear approximation as well as a linear regression model.

PHES Specification		**Day-ahead Market**	
Upper and lower reservoirs		*Energy prices*	[€/MWh]
Shape	Rectangular	Minimum price (4am)	39.9
Bottom surface	$30,000\ m^2$	Maximum price (7pm)	80.3
Total capacity	$735,000\ m^3$	Average price	60.2
Initial volume	$367,500\ m^3$		
Height difference	$74.5\ m$		
Francis turbine		*Reserve capacity prices*	[€/MWh]
Rated power	10 MW	FCR	20
Ramping ability	4 MW/min	aFRR	15
OPEX	3.8 €/MWh	mFRR	10

Table 1: Fictitious PHES and day-ahead market specifications.

The various parameters used to perform the case study are gathered in Table 1. The time-horizon is set to 24 hours (day-ahead scheduling) with an hourly time step. The fictitious PHES plant in Maizeret is made up of two basins separated by a height difference of 74.5 m. Each reservoir is of rectangular shape with a bottom surface of 30,000 m^2 and has a total water capacity of 735,000 m^3. The basins start the day half full (i.e., with 367,500 m^3 of water) while, at the end of the time horizon, the upper reservoir is expected to contain at least 250,000 m^3 of water. The PHES station is equipped with a variable-speed Francis pump-turbine featuring a rated power of 10 MW. Its ramping ability stands at 4 MW/min for both upward and downward directions.

Day-ahead market prices are based on BELPEX data from February 7, 2017, and range from 39.9 €/MWh at 4am to 80.3 €/MWh at 7pm. Figure 6 depicts the positive and negative imbalance prices for the day as retrieved from the Belgian TSO website. The full price profile can be seen in Figure 7. The operating reserve capacities are valued at 20 €/MW for FCR, 15 €/MWfor aFRR and 10 €/MW for mFRR. PHES operating costs are set at 3.8 €/MWh in all modes.

The optimization is solved using Gurobi Optimizer version 10.0.0 build v10.0.0rc2 (mac64[arm]) on an Apple M1 Pro chip with 32GB of RAM. The MIP gap is set at 1% with a time limit of 10 minutes.

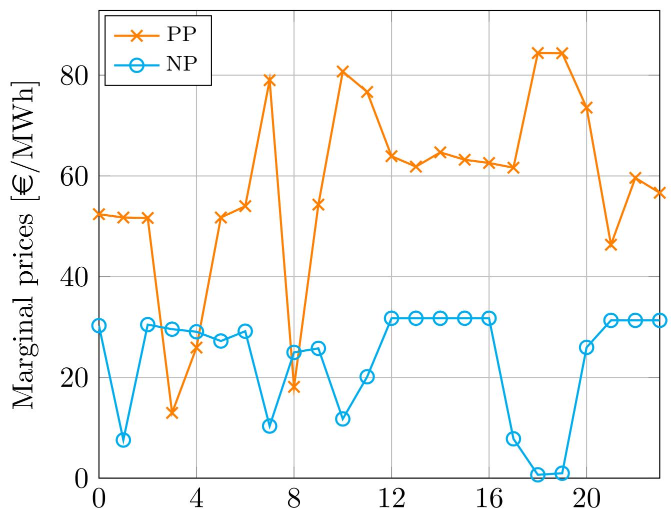

Figure 6: Positive and Negative Prices for February 7, 2017.

4.2. Neural Network Training

The NNs used to model the UPCs are built using Tensorflow. Each NN is trained on a dataset of 50050 data points. Those data points are based on samples collected experimentally on a reduced-scale model of the hydraulic machine. The batch size is set to 16 samples. An early stopping callback is implemented so that training is interrupted when the loss of the validation set is not decreasing for eight consecutive epochs, and the best weights are saved. The quality of the fit, assessed on a test dataset of 500 samples using the coefficient of determination R^2, is presented in Table 2.
Overall, the training time is around one minute, which is insignificant since NNs are trained offline only once before being integrated within the daily optimization routines. Increasing the number of neurons per layer and/or the number of layers positively impacts the accuracy, whether the architecture is sparse or not. Enforcing weight sparsity tends to decrease the resulting NN accuracy, but this effect is marginal. One may observe that the pump UPC is easier to model, which arises from its smoother profile over its smaller operating domain compared to the turbine mode. For equivalent NN architectures, the quality of fit using a single NN is worse than using two. This is particularly true when using a single neuron per layer.

	N° neurons \ N° layers	Conventional 1	Conventional 2	Conventional 3	Sparse 1	Sparse 2	Sparse 3
Turbine	1	0.885	0.891	0.887			
	2	0.985	0.990	0.986	0.953	0.989	0.973
	3	0.996	0.998	0.999	0.987	0.995	0.996
	4	0.997	0.999	0.999	0.996	0.997	0.997
	5	0.998	1.000	0.999	0.994	0.999	0.998
Pump	1	0.813	0.806	0.803			
	2	0.990	0.994	0.991	0.990	0.990	0.991
	3	0.993	0.994	0.999	0.991	0.994	0.993
	4	0.998	0.999	0.999	0.993	0.995	0.998
	5	0.998	0.999	0.999	0.994	0.999	0.999
Both	1	0.225	0.176	0.287			
	2	0.894	0.895	0.941	0.814	0.926	0.888
	3	0.962	0.991	0.994	0.892	0.930	0.958
	4	0.990	0.997	0.999	0.940	0.991	0.997
	5	0.992	0.997	0.999	0.989	0.995	0.998

Table 2: R^2 values of the NNs on the test set at the end of training. Shadow cells highlight the architectures reported in Table 3.

4.3. Model Performances

In this subsection, the impact of different UPC models in the day-ahead PHES scheduling problem is investigated. In particular, both the expected profit (as given at the end of the optimization) and the actual ex-post profit (as computed by the market-based PHES simulator), along with the computation time are analyzed. The best results for each type of model are summarized in Table 3.

NNs are used to approximate the UPCs in two ways: (*i*) each UPC is approximated by its own NN (two NNs are needed in total), and (*ii*) both UPCs are approximated jointly by a single NN. Moreover, an advanced method of Bayesian Optimization (BO), i.e., trust region Bayesian Optimization (TuRBO) [37], was employed. Compared to conventional BO, TuRBO avoids the over-exploration thus offering more resources for exploitation while

[3] The ex-post profit is obtained by subtracting the penalty to the expected profit and accounting for the slight operational cost difference which is not reported in this table.

	N° var.	Solving Time [s]	MIP Gap [%]	Expected [€]	Penalty [€]	Ex-post [€]
2 NN	4710	600	56	2284	-133	*2399*[3]
2NN sp.	3990	594	1	*2321*	-44	2368
1 NN	2694	600	99	2271	-107	2381
1 NN sp.	2694	600	69	2282	-101	2385
sota	3318	27	0	2278	170	2113
l.r.	726	1	0	2399	328	2072
TurBO	30	600	-	-	-	1959

Table 3: Results of (*i*) two conventional NNs, (*ii*) two sparse NNs, (*iii*) one conventional NN, (*iv*) one sparse NN, (*v*) the piecewise state-of-the-art (sota) model with 5 head and 5 power sub-intervals (25 pieces), (*vi*) the linear regression (l.r.) approximation of each UPC and (*vii*) the trust region bayesian optimization algorithm. All the NN models give the best ex-post profit in their respective category.

allowing for heterogeneous modelling. TuRBO relies on a set of local optimization runs. Each local model is robust to noisy samples and estimates precisely the uncertainty, akin to BO. This method calls the simulator to evaluate potential solution in an iterative fashion. Hence, there is no expected profit and the ex-post profit is directly optimized.

The best results in terms of ex-post profit for each type of model are depicted in Table 3, but more detailed outcomes are available in Appendix A. The highest expected profit is given by the sparse 2NN architecture which stands at 2321€. The PHES scheduling obtained with this architecture is displayed in Figure 7.

The plant is operated in turbine mode when the day-ahead prices are high and in pump mode when they are low, thereby making profits through inter-temporal arbitrage in the day-ahead energy market (Figure 7.a). The feasible operational ranges of the hydraulic machine are displayed as light shadow zones and dark shadow zones when accounting for the reserve capacity allocated to the balancing stage. Both the inter-temporal arbitrage and the reserve commitment, which forces the unit to operate in the dark shadow zones, generate revenues.

It can be observed that the power bound approximation is of sufficient quality so that the decisions comply with the feasible power range, except from 6pm to 7pm where the turbine power scheduled is too high, thereby not complying with the upward reserve capacity commitment. The simulator assumes the plant operator does not want to risk any disciplinary issues for

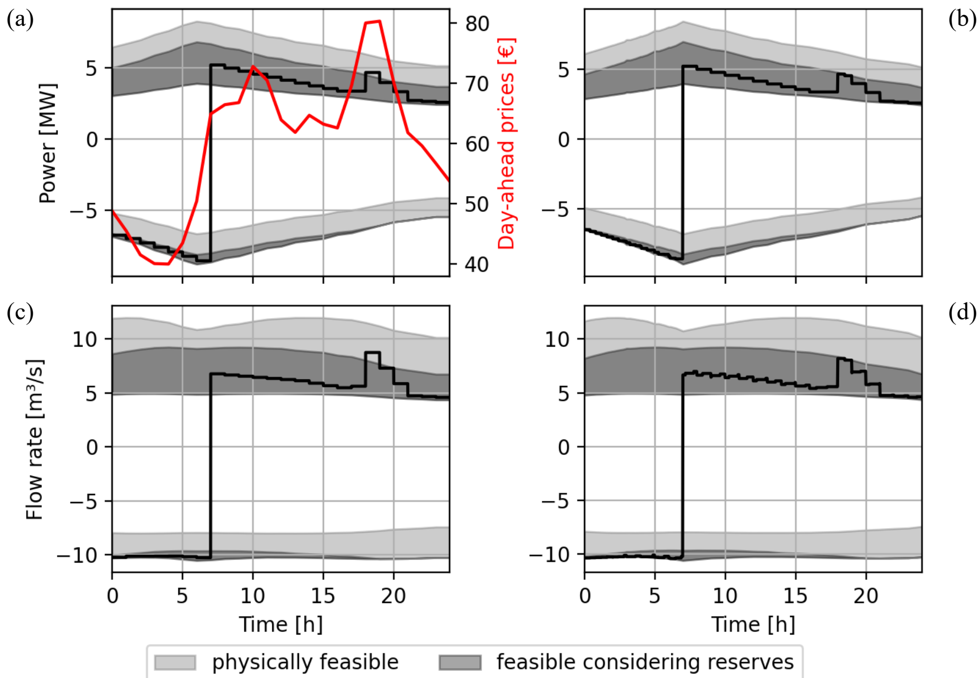

Figure 7: The outcome of the model using two sparse NNs with three layers of four neurons: (a) the electricity prices on the day-ahead market and the power decisions of the optimization, (b) power decisions as corrected by the simulator, (c) the water flow rate decisions of the optimization, (d) water flow rate decisions as corrected by the simulator.

not being able to provide the contracted reserve. Therefore, the simulator continuously adjusts the turbine power over the hour to be as close as possible to its offer on the day-ahead energy market while respecting the reserve commitment (Figure 7.b). It results in a slight penalty of 48€ because of the consequent real time imbalance. Nevertheless, the imbalance penalty is offset by the excess of water remaining in the upper reservoir which adds 91€ leading to an ex-post profit of 2368€. Figure 7.c and Figure 7.d display the water flow rates as decided by the optimization and corrected by the simulator, respectively.

Interestingly, all NN approximations are able to yield higher ex-post profits than the state-of-the-art methods. In particular, using two NNs with three layers of 4 neurons gives the best ex-post profit, 2399€, which is an increase of 11.3% compared to the piecewise linear approach.

For the piecewise linear UPC approximation, the head and the power spaces

are partitioned into five subintervals, which amounts to fit each UPC with 25 planes. The day-ahead schedule is obtained in only 29s for an expected profit of 2278€. The simulator yields an ex-post profit of 2113€ because of 170€ of penalties (see Table 3).
A linear regression model can be used to fit each UPC. Although the resulting problem is still mixed-integer linear due to the remaining presence of binaries to discriminate between operating modes, the solving time plummets to 1s. This is a drastic reduction with respect to the piecewise linear UPC model. The expected profit is higher than for its piecewise counterpart (2399€). However, because the linear approximation is unable to fully capture the complexity of the UPC, high penalties (298€) are incurred, which leads to a smaller ex-post profit (2102€) than the best NN-based formulation, and thus to a high disappointment for the PHES operator. Lastly, the best result obtained with the TuRBO algorithm lies even lower at 1959€. Over the 10 runs, one did not converge. The nine others reaped a average ex-post profit of 1878€. This indicates that the method struggles to capture the complexity of the problem, especially the participation in the reserves due to the high penalties incurred in case of inaccurate reserve commitments.

Figure 8 depicts the average performances of the various NN architectures[4]. Increasing the UPC quality of fit improves the economic value of the PHES scheduling. In particular, for the model with 2 NNs, the average ex-post profit increases from about 2110€ for a neuron to over 2370€ for 5 neurons per layer. It only takes two neurons per layer to outperform the piecewise linear approach. However, when using a single NN, the average ex-post profit starts at 501€ and rises up to 2343€. Such a bad outcome for the one-neuron architectures was to be expected given the poor modeling performances displayed during training. At three neurons per layer, the average ex-post profit stands at 2229€ which is higher than the state-of-the-art. As anticipated, the solving time tends to increase with the number of neurons per layer. It should be noted that some NN architectures with a higher number of layers and neurons involve a computation time higher than the target time limit of 10 minutes. For these cases, the MIP gap is reported.

Interestingly, the approach with two NNs outperforms its single NN coun-

[4]Since the weight pruning rate is set to 25%, no pruning is conducted for layers containing less than four weights. Therefore, the sparse architectures with hidden layers of one neuron are not reported since they are equivalent to the ones without sparsity.

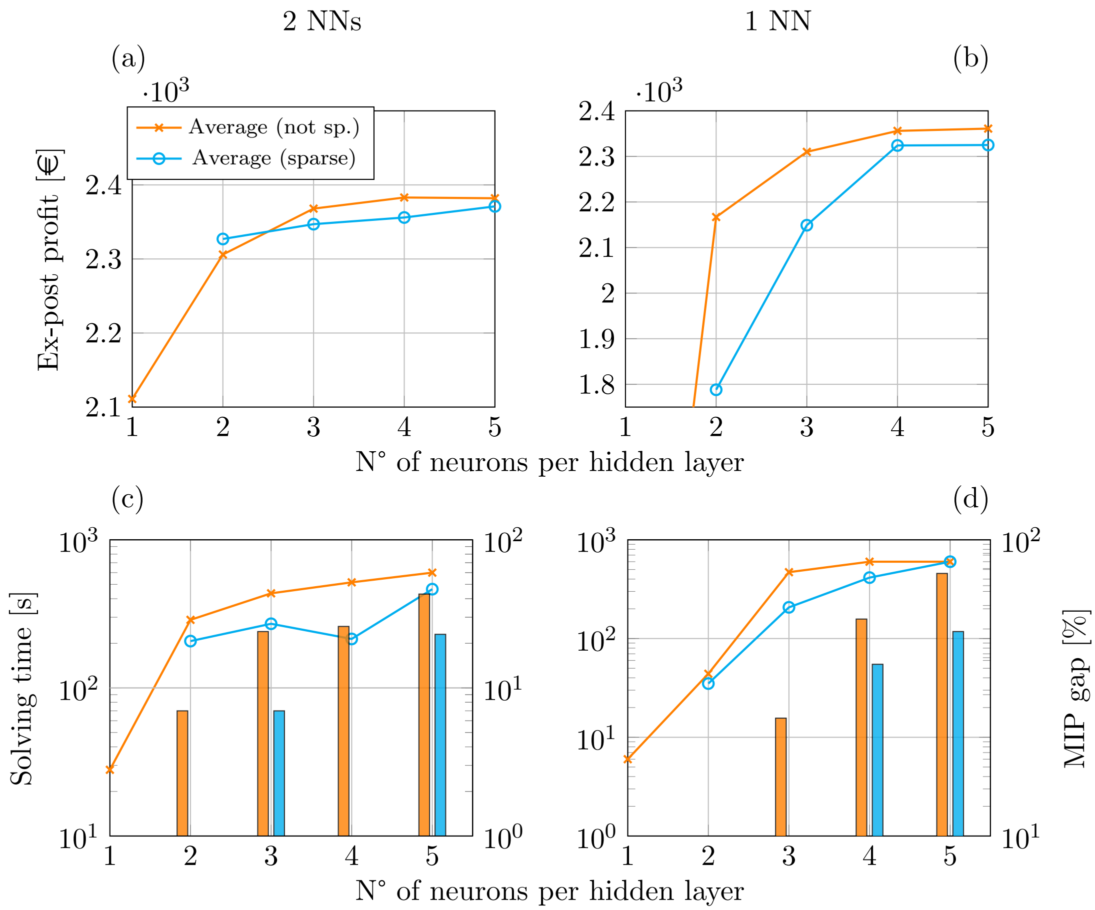

Figure 8: Ex-post profits obtained using (left-hand side) two NNs (one per UPC); (right-hand side) a single NN for both UPCs.

terpart in terms of *average* ex-post profit. Moreover, the former is faster to solve for more complex architectures and exhibits tighter MIP gaps after 10 minutes.

4.3.1. The impact of weight pruning

The sparse architectures yield slightly lower ex-post profits, at the exception of the architectures using a single NN with four neurons per layer (Figure 8). Since the pruning adds constraints to the NN training, it reduces the modeling power of the resulting model. Still, such sparse NNs manage to be very competitive and sometimes even outperform their conventional counterpart, i.e., the best outcome for the single NN model is provided by

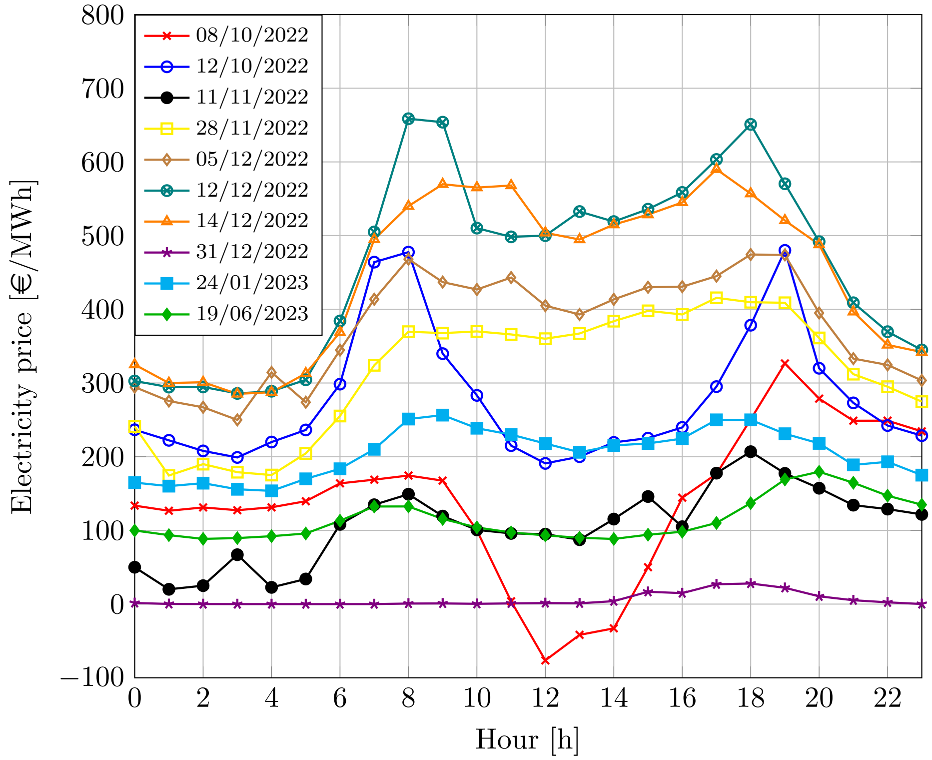

Figure 9: Day-ahead electricity prices in Belgium for the ten days selected.

a sparse architecture, as reported in Table 3. Moreover, sparse architectures tend to be quicker to solve, and if they are not solved in 10 minutes, the MIP gap is smaller.

4.4. Performances over multiple scenarios

In order to assess the performance of the proposed methodology over a representative number of scenarios, 10 days spanning from December 2022 to September 2023 exhibiting different price levels and dynamics were selected. For each day, the day-ahead electricity price were retrieved from Belpex. Figure 9 depicts the price profile for these days. The impact of the water volume in the upper reservoir at the beginning of each day is also evaluated. Three configurations are studied wherein the upper reservoir is respectively filled to 40%, 50% and 60%. In total, thirty scenarios are thus obtained, three for each of the ten days.

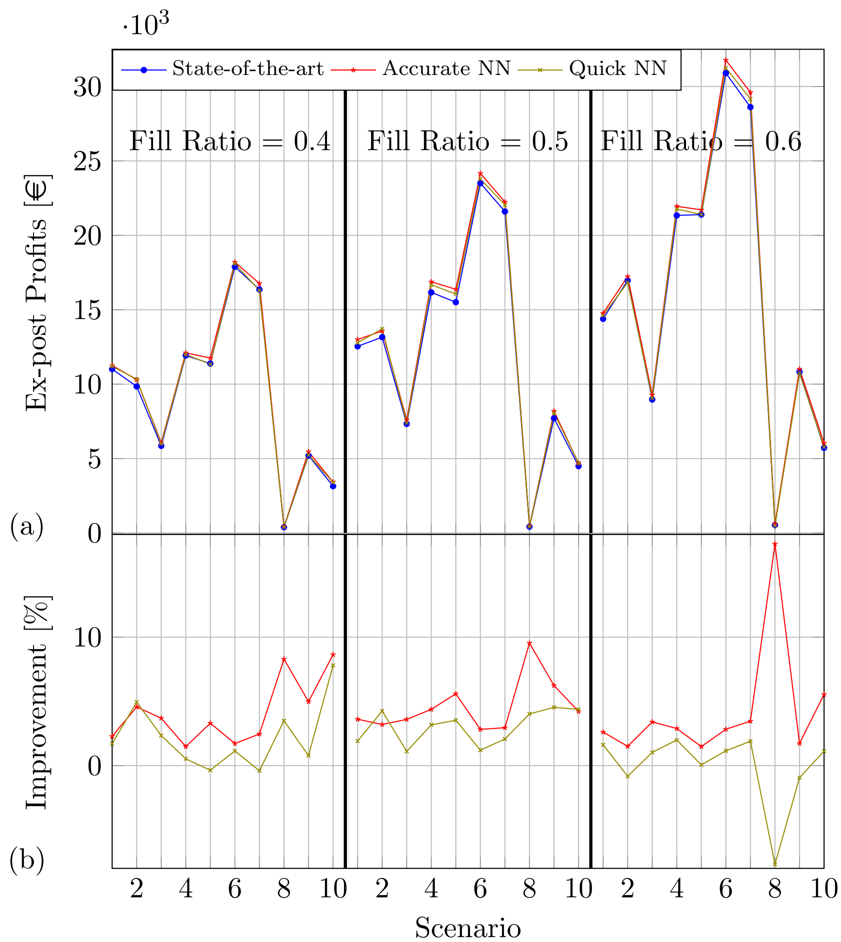

Figure 10: (a) Ex-post profit [€] of each scenario for the state-of-the-art piecewise method with 25 planes, the 2NN-approximation with 3 hidden layers of 4 neurons and the 2NN-approximation with one sparse hidden layers of 4 neurons; (b) Improvement [%] of the NN models with respect to the state-of-the-art model.

Three models have been selected based on the outcomes of subsection 4.3: (i) the best state-of-the-art model featuring 25 pieces, (ii) one model using two sparse NNs of one hidden layer with four neurons and (iii) one model using one sparse NN with three hidden layers of five neurons. The second model is chosen for its trade-off between accuracy and computation burden while the third model delivered the best ex-post profit on average.

The ex-post profit for each of the thirty scenarios is reported in Figure 10 (solid lines). The NN-based optimization is chosen for its accuracy that consistently outperforms the state-of-the-art. The average ex-post profit achieved by the accurate model is 12,897€, 4.3% higher than the state-of-the-art at 12,504€ while the second NN model reaches 12,702€ of average ex-post profit (i.e., 1.6% improvement).

The relative improvements in % brought by the NN models with respect to the state-of-the-art model are also depicted in Figure 10 (dashed lines). It can be seen that the Accurate NN model consistently outperforms the state-of-the-art model. It also exhibits higher gains that the Quick NN model at the exception of price profile 2 for fill ratios at 0.4 and 0.5.

Regarding solving times, the quickest model is the state-of-the-art that provides results in 24s on average. Then, the quick NN model is solved in 106s, while the Accurate NN takes 440s on average. Therefore, the quite consistent gain in profitability comes at the cost of a longer solving time. Interestingly, the NN architecture can be chosen to obtain the desired trade-off between accuracy (and thus higher ex-post profits) and solving time.

4.5. Sensitivity analysis to real-time penalties

The sensitivity of ex-post profits to penalties arising from the inability to comply with the optimized day-ahead schedule is evaluated on the same scenario as in Subsection 4.3, i.e., the price signal is from February 7, 2017 with a starting upper reservoir half full. The same three models as in Subsection 4.4 are used.

There are four types of penalties:

1. The penalty for each MW of reserve capacity which cannot be provided. A commitment to the reserve must be maintained; otherwise, the operator might be kicked out of the reserve market by the TSO. Therefore, the minimum penalty is set to 500 €/MW for and can go up to 2000 €/MW for each hour of overcommitted reserve capacity. The range is covered by steps of 100 €/MW.

2. The imbalance settlement price which is incurred to the PHES operator for deviating from their energy bid in the day-ahead market. The imbalance settlement costs, namely PP and NP, are retrieved for each day from the Belgian TSO's website.

3. The cost associated with the lack of water in the upper basin at the end of the time horizon. This water represents energy which cannot be generated in the future because it was used in excess on the present

day. The excess of expected profit by over-using the water can be evaluated between the minimum and the maximum electricity prices over the present day. This range is covered by steps of 5 €/MWh.

4. The value associated with the excess of water in the upper basin at the end of the time horizon. This excess of water represents energy which could be traded on the following days. One extreme assumption is to consider this water valueless. On the contrary, one might argue it could be sold the following day at the maximum day-ahead price. Hence, the excess of water is considered to vary between 0 €/MWh and the maximum electricity price on the given day, with steps of 5 €/MWh.

Table 4 summarizes the possible range for each penalty. In total, the impact of 2484 combinations of parameters on the results of three models are studied.

	Values	**Unit**
Reserve Overcommitment	500 : 100 : 2000	€/MW
Imbalance Settlement	TSO's	€/MWh
Lack of water	$\lambda_t^{\mathrm{DA,min}}$: 5 : $\lambda_t^{\mathrm{DA,max}}$	€/MWh
Excess of water	0 : 5 : $\lambda_t^{\mathrm{DA,max}}$	€/MWh

Table 4: Description of the various penalty values for each type of deviation studied in the sensitivity analysis.

For each model, a normal distribution of the ex-post profits are obtained and represented in Figure 11. Firstly, one can see that the distributions are quite compact with standard deviations at 35.1€, 25€ and 12.9€ for the state-of-the-art, the Accurate NN and the Quick NN models, respectively. It means the ex-post results reported do not vary significantly with respect to the penalty values. Secondly, the average profit of the state-of-the-art model stands at 2119€ well under 2313€ and 2309€ for the quick and accurate NNs, respectively. Surprisingly, *under the proposed sampling*, the Quick NN model results are very slightly higher on average and more concentrated than the ones of the Accurate NN model. Thus, the Quick NN model is more robust to changes in penalty values.

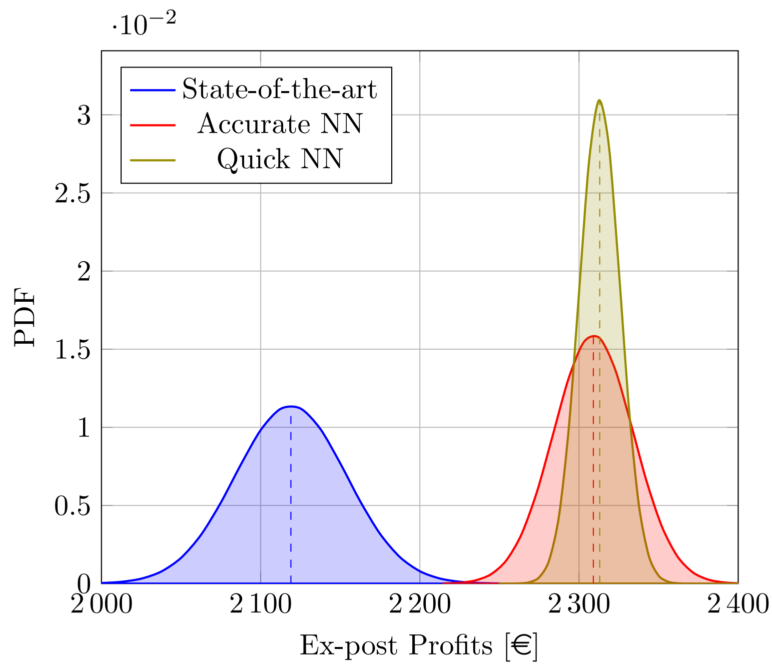

Figure 11: Normal distribution of the ex-post profits sensitivity to the penalties.

5. Conclusion

In this work, NN-constrained optimization is introduced to improve the quality of the PHES day-ahead scheduling problem by enhancing the accuracy of the UPCs modeling. Two approaches are developed. The first one replaces each of the UPC (in turbine and pump modes) by its NN approximation. The second relies on a single NN to approximate both curves. The NNs are reformulated exactly as a set of MILP constraints and embedded into the PHES optimization in day-ahead and reserve markets.

In the case study, the best performance is achieved by the proposed NN-constrained optimization, considering a fictitious PHES plant on a Belgian site. Over thirty scenarios, the best NN model was able to improve the ex-post profit of the state-of-the-art by 4.3%. The validity of the results for different ex-post penalty values was confirmed by looking at 2484 combinations. The solving time of the optimization problem tends to quickly increase with the complexity of the NN architecture. Pruning weights alleviates this problem by reducing the computational burden with a limited loss in the quality of the decisions. Such sparse architectures are thus natural candidates for capturing non-linear relationships in optimization problems. Overall, if the focus is to reach the optimal decisions, NN-informed optimization is a great tool for models presenting multidimensional non-linear relationships. Moreover, the architecture of the NNs can be adapted to find the right trade-off between accuracy and computational burden. In the present case study, results support the use of one NN per non-linear relationship. If the solving time must be under one minute, state-of-the-art piecewise approximations perform well. For close to real-time decision-making, e.g., with a one second time resolution, at the expense of the decision quality, a linear regression of the UPCs is a promising lead. Bayesian optimization algorithms are not able to outperform formal optimization and the randomness of the results is another strong barrier their adoption.

This data-driven approach is generic and can be used in any optimization problem wherein one or several complex, physically unknown, relationships need to be modelled in a tractable fashion using the robustness of MILP solvers. The findings of this study offer valuable insights into the practical applications of neural network-constrained optimization by addressing the challenges of PHES modeling. The work can be extended to capture the market price uncertainty. Furthermore, the PHES operational cost model could more accurately account for the wear and tear of the hydraulic machine. Lastly, it would be interesting to investigate whether other NN reformulations (e.g., using tighter equations for the ReLU function) can help decrease the computation time.

Acknowledgement

This work was funded by the F.R.S.-FNRS under the grant number *FC 49537*.

The authors wish to thank deeply Maxime Gobert and Adriano Arrigo for their help and advice.

Appendix A: Results of the NN informed models

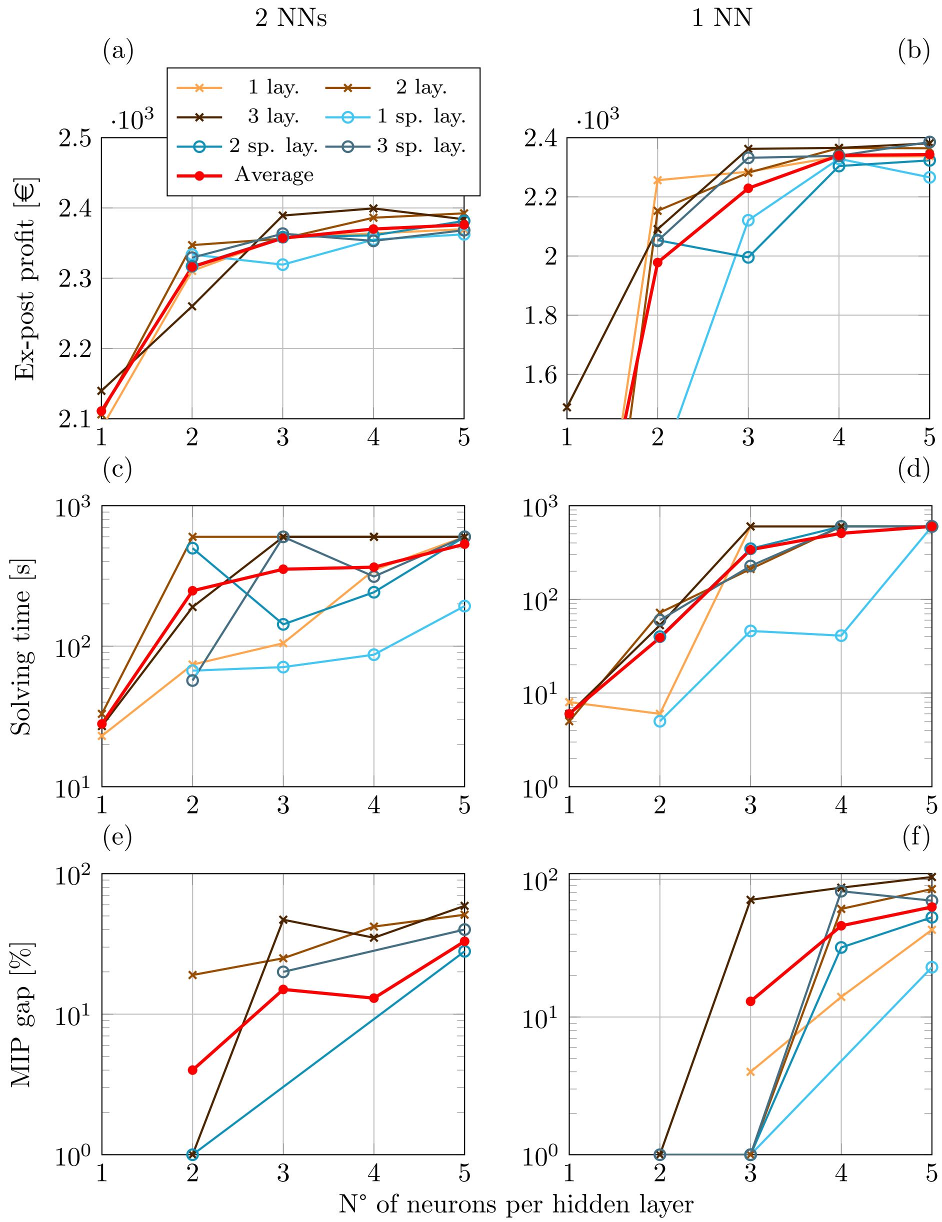

Figure A.1: Ex-post profits obtained using (left-hand side) two NNs (one per UPC); (right-hand side) a single NN for both UPCs.

Sparse	**N° layers**	**N° neurons**	**N° var.**	**Solv. Time[s]**	**MILP gap [%]**	**Expected [€]**	**MAE on Q**	**ME on Q**	**Penalty [€]**	**Ex-post [€]**
No	1	1	1350	23	0.01	2548.3	0.34	0.219	468.39	2086.3
No	1	2	1590	74	0	2323.8	0.19	-0.041	17.14	2310.4
No	1	3	1830	105	0	2346	0.16	-0.066	-8.12	2357.9
No	1	4	2070	348	0	2286.1	0.16	-0.12	-74.61	2363.7
No	1	5	2310	600	0.2	2265.1	0.15	-0.143	-101.76	2369.8
No	2	1	1590	33	0	2532.3	0.32	0.198	432.45	2106.2
No	2	2	2070	600	0.19	2310.8	0.16	-0.089	-32.6	2347.1
No	2	3	2550	600	0.25	2304.8	0.11	-0.09	-48.47	2356.3
No	2	4	3030	600	0.42	2274.6	0.15	-0.148	-108.88	2386.1
No	2	5	3510	600	0.51	2270.3	0.16	-0.161	-119.25	2392.3
No	3	1	1830	27	0	2592.5	0.26	0.211	459.11	2139.7
No	3	2	2550	190	0.01	2369.1	0.19	0.042	111.77	2260
No	3	3	3270	600	0.47	2276.3	0.16	-0.155	-109.99	2389.2
No	3	4	3990	600	0.35	2264	0.17	-0.174	-132.68	2399.3
No	3	5	4710	600	0.59	2283.5	0.14	-0.136	-97.86	2383.9
Yes	1	1	1350	19	0.01	2541.2	0.34	0.209	451.46	2096.2
Yes	1	2	1590	67	0	2317.4	0.24	-0.092	-10.56	2333.7
Yes	1	3	1830	71	0	2301.7	0.18	-0.066	-14.09	2319.3
Yes	1	4	2070	87	0	2363.3	0.16	-0.049	12.32	2354.9
Yes	1	5	2310	193	0	2339.1	0.16	-0.074	-19.74	2362.4
Yes	2	1	1590	18	0.01	2551.4	0.33	0.222	474.35	2083.5
Yes	2	2	2070	498	0.01	2309.7	0.18	-0.058	-3.64	2317
Yes	2	3	2550	143	0	2318	0.21	-0.092	-37.11	2358.4
Yes	2	4	3030	242	0	2301.3	0.13	-0.098	-56.85	2360.9
Yes	2	5	3510	600	0.28	2282.5	0.14	-0.139	-96.31	2381.5
Yes	3	1	1830	3	0	2527.7	0.21	0.135	318.65	2214.2
Yes	3	2	2550	57	0	2329.6	0.26	-0.078	6.26	2329.2
Yes	3	3	3270	600	0.2	2277.9	0.16	-0.128	-82.61	2363.4
Yes	3	4	3990	313	0	2314.2	0.12	-0.08	-35.99	2353.1
Yes	3	5	4710	600	0.4	2321	0.12	-0.092	-44.36	2368.3

Table A.1: Result of the optimization models using one NN for modelling each UPCs (2 NNs in total).

Sparse	**N° layers**	**N° neurons**	**N° var.**	**Solv. Time[s]**	**MILP gap [%]**	**Expected [€]**	**MAE on Q**	**ME on Q**	**Penalty [€]**	**Ex-post [€]**
No	1	1	1014	8	0	3540.6	1.65	1.61	3413.0	181.9
No	1	2	1134	6	0	2476.1	0.19	0.061	228.1	2256.3
No	1	3	1254	600	0.04	2356.8	0.22	-0.001	77.5	2284.9
No	1	4	1374	600	0.14	2380.7	0.3	-0.045	50.2	2336.5
No	1	5	1494	600	0.43	2366.3	0.29	-0.052	34.5	2337.2
No	2	1	1134	5	0	3783.2	1.96	1.848	4008.5	-167.7
No	2	2	1374	72	0	2586	0.26	0.198	439.6	2152.9
No	2	3	1614	211	0.01	2339.2	0.14	0.01	60.7	2281.6
No	2	4	1854	600	0.61	2310.8	0.2	-0.112	-50.5	2365.1
No	2	5	2094	600	0.85	2324.5	0.21	-0.1	-35.9	2364.3
No	3	1	1254	6	0	1861.7	0.9	0.107	381.6	1488.5
No	3	2	1614	53	0.01	2579.4	0.35	0.239	493.9	2090.7
No	3	3	1974	600	0.71	2321.5	0.2	-0.099	-36.8	2362.2
No	3	4	2334	600	0.87	2274.3	0.14	-0.134	-88.1	2365.5
No	3	5	2694	600	1.04	2270.8	0.15	-0.147	-107.4	2380.9
Yes	1	1	1014	8	0.01	3544.9	1.67	1.632	3417.5	182
Yes	1	2	1134	5	0	2511	0.63	0.602	1264.3	1260.6
Yes	1	3	1254	46	0.01	2525.2	0.35	0.187	410.4	2120.8
Yes	1	4	1374	41	0	2301.3	0.27	-0.105	-21.4	2328.7
Yes	1	5	1494	600	0.23	2446.9	0.43	0.036	189.6	2265.9
Yes	2	1	1134	5	0	3766.1	1.96	1.842	3993.3	-169.9
Yes	2	2	1374	40	0	2611.7	0.52	0.277	563.8	2053.5
Yes	2	3	1614	347	0.01	2636.8	0.47	0.342	645.4	1995.5
Yes	2	4	1854	600	0.32	2372.6	0.18	0.001	73.5	2304.1
Yes	2	5	2094	600	0.53	2322.6	0.13	-0.056	3.8	2323.3
Yes	3	1	1254	15	0	3540.6	1.65	1.61	3412.0	182.9
Yes	3	2	1614	60	0.01	2568.3	0.37	0.263	522.3	2051.3
Yes	3	3	1974	227	0.01	2358.4	0.29	-0.058	32.4	2331.9
Yes	3	4	2334	600	0.82	2309.4	0.09	-0.061	-26.8	2338.8
Yes	3	5	2694	600	0.7	2282	0.14	-0.134	-101.1	2385.3

Table A.2: Result of the optimization models using a single NN for modelling both UPCs.

Appendix B: Results of the TuRBO algorithm

Run N°	**Ex-post Profit [€]**
1	1887.79035
2	1832.12974
3	1862.14761
4	1854.69737
5	-3252.7209
6	1898.24603
7	1832.82968
8	1907.62627
9	**1959.44507**
10	1867.20086

Table A.3: Results of the TuRBO algorithm over the scenario presented in Section 4.3.

References

[1] European Commission, Directorate-General for Energy, C. Andrey, P. Barberi, L. Nuffel, F. Gérard, J. Gorenstein Dedecca, K. Rademaekers, Y. El Idrissi, M. Crenes, L. Lacombe, Study on energy storage : contribution to the security of the electricity supply in Europe, Publications Office, 2020. doi:10.2833/077257.

[2] DOE Global Energy Storage Database - Statistics (2022).
URL https://sandia.gov/ess-ssl/gesdb/public/statistics.html

[3] S. Nag, K. Y. Lee, A Comparison of the Dynamic Performance of Conventional and Adjustable-Speed Pumped Storage Hydro, in: 2021 IEEE Power & Energy Society General Meeting (PESGM), 2021, pp. 01–05, journal Abbreviation: 2021 IEEE Power & Energy Society General Meeting (PESGM). doi:10.1109/PESGM46819.2021.9638132.

[4] A. Morabito, P. Hendrick, Pump as turbine applied to micro energy storage and smart water grids: A case study, Applied Energy 241 (2019) 567–579. doi:10.1016/j.apenergy.2019.03.018.
URL https://www.sciencedirect.com/science/article/pii/S030626191930409X

[5] E. Barbour, I. G. Wilson, J. Radcliffe, Y. Ding, Y. Li, A review of pumped hydro energy storage development in significant international electricity markets, Renewable and Sustainable Energy Reviews 61 (2016) 421–432, publisher: Elsevier Ltd.
URL https://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S1364032116300363&lang=fr&site=eds-live

[6] E. Nobile, G. Sari, A. Schwery, Variable Speed Hydro Pumped Storage as Flexible Enabler of Intermittent Renewable Penetration, in: 2018 IEEE Power & Energy Society General Meeting (PESGM), 2018, pp. 1–5, journal Abbreviation: 2018 IEEE Power & Energy Society General Meeting (PESGM). doi:10.1109/PESGM.2018.8586238.

[7] Ø. S. Klyve, V. J. Olkkonen, A. Hensel, Techno-economic viability of pumped hydropower plants in Germany, with focus on battery hybridization, in: 7th International Hybrid Power Plants & Systems Workshop (HYB 2023), Vol. 2023, 2023, pp. 74–82, journal Abbreviation: 7th

International Hybrid Power Plants & Systems Workshop (HYB 2023). doi:10.1049/icp.2023.1435.

[8] T. Mercier, J. Jomaux, E. De Jaeger, M. Olivier, Provision of primary frequency control with variable-speed pumped-storage hydropower, in: 2017 IEEE Manchester PowerTech, 2017, pp. 1–6, journal Abbreviation: 2017 IEEE Manchester PowerTech. doi:10.1109/PTC.2017.7980920.

[9] SMARTWATER (Sep. 2022).
URL https://www.multitel.be/projets/smartwater/

[10] Y. Pannatier, Optimisation des stratégies de réglage d'une installation de pompage-turbinage à vitesse variable - Infoscience, Ph.D. thesis, Université de Lausanne (Jan. 2018).
URL https://infoscience.epfl.ch/record/149805/files/

[11] M. Daadaa, S. Séguin, K. Demeester, M. F. Anjos, An optimization model to maximize energy generation in short-term hydropower unit commitment using efficiency points, International Journal of Electrical Power & Energy Systems 125 (2021) 106419. doi:10.1016/j.ijepes.2020.106419.
URL https://www.sciencedirect.com/science/article/pii/S0142061519342218

[12] C. Ju, T. Ding, W. Jia, C. Mu, H. Zhang, Y. Sun, Two-stage robust unit commitment with the cascade hydropower stations retrofitted with pump stations, Applied Energy 334 (2023) 120675. doi:10.1016/j.apenergy.2023.120675.
URL https://www.sciencedirect.com/science/article/pii/S0306261923000399

[13] M. Zhang, J. Chen, Z. Yang, K. Peng, Y. Zhao, X. Zhang, Stochastic day-ahead scheduling of irrigation system integrated agricultural microgrid with pumped storage and uncertain wind power, Energy 237 (2021) 121638. doi:10.1016/j.energy.2021.121638.
URL https://www.sciencedirect.com/science/article/pii/S0360544221018867

[14] J.-F. Toubeau, Z. De Greve, P. Goderniaux, F. Vallee, K. Bruninx, Chance-Constrained Scheduling of Underground Pumped Hydro

Energy Storage in Presence of Model Uncertainties, IEEE transactions on sustainable energy 11 (3) (2020) 1516–1527, number: 3 Place: PISCATAWAY Publisher: PISCATAWAY: IEEE. doi:10.1109/TSTE.2019.2929687.

[15] J. Jamii, M. Trabelsi, M. Mansouri, M. F. Mimouni, W. Shatanawi, Non-Linear Programming-Based Energy Management for a Wind Farm Coupled with Pumped Hydro Storage System, Sustainability 14 (18) (2022). doi:10.3390/su141811287.

[16] J. Catalão, H. Pousinho, J. Contreras, Optimal hydro scheduling and offering strategies considering price uncertainty and risk management, Energy (Oxford) 37 (1) (2012) 237–244, place: OXFORD Publisher: Elsevier Ltd. doi:10.1016/j.energy.2011.11.041.

[17] E. C. Finardi, E. L. da Silva, Solving the hydro unit commitment problem via dual decomposition and sequential quadratic programming, IEEE Transactions on Power Systems 21 (2) (2006) 835–844. doi:10.1109/TPWRS.2006.873121.

[18] N. Mousavi, G. Kothapalli, D. Habibi, C. K. Das, A. Baniasadi, A novel photovoltaic-pumped hydro storage microgrid applicable to rural areas, Applied Energy 262 (2020) 114284. doi:10.1016/j.apenergy.2019.114284.
URL https://www.sciencedirect.com/science/article/pii/S0306261919319713

[19] Y. Chen, M. Wang, Y. Zhang, Y. Lu, B. Xu, L. Yu, Cascade Hydropower System Operation Considering Ecological Flow Based on Different Multi-Objective Genetic Algorithms, Water Resources Management 37 (8) (2023) 3093–3110. doi:10.1007/s11269-023-03491-3.
URL https://doi.org/10.1007/s11269-023-03491-3

[20] Y. H. Al-Aqeeli, O. M. A. Mahmood Agha, Optimal Operation of Multi-reservoir System for Hydropower Production Using Particle Swarm Optimization Algorithm, Water Resources Management 34 (10) (2020) 3099–3112. doi:10.1007/s11269-020-02583-8.
URL https://doi.org/10.1007/s11269-020-02583-8

[21] L. E. Schäffer, A. Helseth, M. Korpås, A stochastic dynamic programming model for hydropower scheduling with state-dependent maximum discharge constraints, Renewable Energy 194 (2022) 571–581. doi:10.1016/j.renene.2022.05.106.
URL https://www.sciencedirect.com/science/article/pii/S0960148122007534

[22] T. Zhao, J. Zhao, D. Yang, Improved Dynamic Programming for Hydropower Reservoir Operation, Journal of Water Resources Planning and Management 140 (3) (2014) 365–374, publisher: American Society of Civil Engineers. doi:10.1061/(ASCE)WR.1943-5452.0000343.
URL https://ascelibrary.org/doi/10.1061/%28ASCE%29WR.1943-5452.0000343

[23] E. Finardi, F. Takigawa, B. Brito, Assessing solution quality and computational performance in the hydro unit commitment problem considering different mathematical programming approaches, Electric Power Systems Research 136 (2016) 212–222. doi:10.1016/j.epsr.2016.02.018.
URL https://www.sciencedirect.com/science/article/pii/S0378779616300244

[24] X. Cheng, S. Feng, H. Zheng, J. Wang, S. Liu, A hierarchical model in short-term hydro scheduling with unit commitment and head-dependency, Energy (Oxford) 251 (2022) 123908, place: OXFORD Publisher: Elsevier Ltd. doi:10.1016/j.energy.2022.123908.

[25] C. Cheng, C. Su, P. Wang, J. Shen, J. Lu, X. Wu, An MILP-based model for short-term peak shaving operation of pumped-storage hydropower plants serving multiple power grids, Energy 163 (2018) 722–733. doi:10.1016/j.energy.2018.08.077.
URL https://www.sciencedirect.com/science/article/pii/S0360544218316098

[26] G. E. Alvarez, Operation of pumped storage hydropower plants through optimization for power systems, Energy 202 (2020) 117797. doi:10.1016/j.energy.2020.117797.
URL https://www.sciencedirect.com/science/article/pii/S036054422030904X

[27] H. Gomes e Souza, E. C. Finardi, B. H. Brito, F. Y. K. Takigawa, Partitioning approach based on convex hull and multiple choice for solving hydro unit-commitment problems, Electric Power Systems Research 211 (2022) 108285. doi:10.1016/j.epsr.2022.108285.
URL https://www.sciencedirect.com/science/article/pii/S0378779622004783

[28] B. Tong, Q. Zhai, X. Guan, An MILP Based Formulation for Short-Term Hydro Generation Scheduling With Analysis of the Linearization Effects on Solution Feasibility, IEEE transactions on power systems 28 (4) (2013) 3588–3599, number: 4 Place: PISCATAWAY Publisher: PISCATAWAY: IEEE. doi:10.1109/TPWRS.2013.2274286.

[29] A. Kody, S. Chevalier, S. Chatzivasileiadis, D. Molzahn, Modeling the AC power flow equations with optimally compact neural networks: Application to unit commitment, Electric Power Systems Research 213, publisher: Elsevier B.V. (Dec. 2022).
URL https://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S0378779622004771&lang=fr&site=eds-live

[30] S. Casarin, G. Cavazzini, J. I. Pérez-Díaz, Battery and Flywheel hybridization of a reversible Pumped-Storage Hydro Power Plant for wear and tear reduction, Journal of energy storage 71 (2023) 108059, place: AMSTERDAM Publisher: Elsevier Ltd. doi:10.1016/j.est.2023.108059.

[31] A. Arce, T. Ohishi, S. Soares, Optimal Dispatch of Generating Units of the Itaipu Hydroelectric Plant, Power Systems, IEEE Transactions on 17 (2002) 154–158. doi:10.1109/59.982207.

[32] S. A. BILLINGS, H. B. JAMALUDDIN, S. CHEN, Properties of neural networks with applications to modelling non-linear dynamical systems, International Journal of Control 55 (1) (1992) 193–224, publisher: Taylor & Francis. doi:10.1080/00207179208934232.
URL https://doi.org/10.1080/00207179208934232

[33] J. Suykens, J. Vandewalle, B. de Moor, Artificial Neural Networks for Modelling and Control of Non-Linear Systems, Springer US, 2012.
URL https://books.google.co.uk/books?id=tmTTBwAAQBAJ

[34] C. Baldassi, E. M. Malatesta, R. Zecchina, Properties of the Geometry of Solutions and Capacity of Multilayer Neural Networks with Rectified Linear Unit Activations, Physical review letters 123 (17) (2019) 1–170602, place: COLLEGE PK Publisher: Amer Physical Soc. doi:10.1103/PhysRevLett.123.170602.

[35] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, Journal of machine learning research 15 (2011) 315–323.

[36] I. Murzakhanov, A. Venzke, G. S. Misyris, S. Chatzivasileiadis, Neural Networks for Encoding Dynamic Security-Constrained Optimal Power Flow, arXiv:2003.07939 [cs, eess, math]ArXiv: 2003.07939 (Oct. 2021). URL http://arxiv.org/abs/2003.07939

[37] D. Eriksson, M. Pearce, J. R. Gardner, R. Turner, M. Poloczek, Scalable Global Optimization via Local Bayesian Optimization, arXiv:1910.01739 [cs, stat] (Feb. 2020).
URL http://arxiv.org/abs/1910.01739