

Affine Policies and Principal Components Analysis for Self-Scheduling in CAES Facilities

Matheus F. Zambroni de Souza, *Student Member, IEEE*, Claudio A. Cañizares, *Fellow, IEEE*, Kankar Bhattacharya, *Fellow, IEEE*, and Alvaro Lorca, *Member, IEEE*

Abstract—This paper presents a novel methodology based on Principal Components Analysis (PCA) and Affine Policies (AP) for self-scheduling of a price-taker Compressed Air Energy Storage (CAES) facility operating under uncertainties. The proposed PCA-AP model is developed from the facility owner's perspective, which partakes in energy, spinning, and idle reserve markets. A methodology is proposed to select the required price uncertainty intervals from actual data based on a Box Cox technique. For a more realistic representation, the detailed thermodynamic characteristics of the CAES facility are considered, taking into account as well modern CAES facilities that may charge and discharge concurrently. To validate the proposed PCA-AP model and approach, the results obtained are compared with an existing Affine Arithmetic (AA) model, which is also based on an affine approach, and Monte Carlo Simulations (MCS), which can be considered as the benchmark for comparison purposes. The input data, forecast prices and intervals of uncertainty, are taken from the Ontario-Canada electricity market for 2015-2019. From the studies presented, it can be observed that the new PCA-AP approach provides less conservative results as compared to the AA approach, and hence can be considered an adequate methodology for day-ahead operations in systems with significant sources of uncertainty.

Index Terms—Affine Arithmetic (AA), Affine Policies (AP), Compressed Air Energy Storage (CAES), price uncertainties, Principal Components Analysis (PCA), self-scheduling.

NOMENCLATURE

Indices

0	Center value.
h	Noise terms.
n	Principal Component from 1 to NPC.
s	Segment.
t	Operation intervals from 1 to T .

Parameters

ang_s^D	Slope of the Discharged Power function at segment s .
$\bar{b}^D$	Length of each segment (MW).
$\overline{CA}$	Max. mass of air the cavern can store (kg).

This work was supported by Canada's NSERC Energy Storage Technology (NEST) Network (<https://nrc.canada.ca/en>).

This work was supported by Canada's NSERC Energy Storage Technology (NEST) Network.

M. F. Z. de Souza, C. A. Cañizares and K. Bhattacharya are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: m6ferrei@uwaterloo.ca; ccanizar@uwaterloo.ca; kankar@uwaterloo.ca).

A. Lorca is with the Department of Electrical Engineering, the Department of Industrial and Systems Engineering, and the UC Energy Research Center at Pontificia Universidad Católica de Chile, Santiago, Chile (e-mail: alvarolorca@uc.cl).

$\overline{\Delta\pi_t}$
 $\overline{\Delta PC_n}$

$\frac{D_s}{\Lambda^Y}$
 $\mathbf{V}^Y$
 ε^Y
 M
 $\overline{NG}, \overline{NG}$

NPC
 PC_n
 π_t^Y

Π^Y
 π^{NG}
 $\underline{P}^X, \overline{P}^X$

q_s
 $\overline{QSC}$
 $\overline{SOC}, \overline{SOC}$

SOC^f
 $VOM^{c/e}$

Variables

A
 $air^{C/D}$
 B
 b_s^D
 CO^{NG}
 $\Delta air^{C/D}$
 ΔCO^{NG}
 ΔP^X
 ΔPC_n
 $\Delta \pi^Y$
 $\Delta \pi^{Y+/-}$
 ΔSOC
 Δw
 OC
 P^X
 $P_{n,t}^{X,Y}$

Max. price mismatch at t (%).

Max. uncertainty associated with Principal Component n .

Min. air discharged for segment s (kg/s).

Eigenvalue matrix for electricity price Y .

Eigenvector matrix for electricity price Y .

Noise term of Y .

Large number, for the Big M method.

Min. and max. heat rate of the facility (MBTU/h).

Number of Principal Components.

Principal Component n .

Price of Y , i.e., energy E , spinning SR and idle ID reserve, at time t (\$/MWh).

Matrix for electricity price Y .

Price of natural gas (\$/MBTU).

Min. and max. limits of X mode, i.e., discharging D , charging C , spinning reserve discharging SRD , spinning reserve charging SRC or idle ID of the CAES facility (MW).

Min. power discharged for segment s (MW).

Quick start capacity of the facility (MW).

Min. and max. limits for the State Of Charge (%).

Final State of Charge (%).

Variable operation and maintenance cost of the compressor c or expander e (\$/MWh).

Revenue from energy arbitrage (\$/h).

Amount of air charged/discharged (kg/s).

Revenue from spinning reserve (\$/h).

Fractional value of segment s (MW).

Cost of natural gas (\$/h).

Deviation of air charged/discharged (kg/s).

Deviation of cost of natural gas (\$/h).

Deviation of P^X (MW).

Deviation of PC n .

Price deviation of Y (%).

Upward/Downward price deviation of Y (%).

Deviation of State of Charge (%).

Deviation of w .

Operation cost (\$/h).

Power in X mode (MW).

Adjustment for power dispatch X , for price Y , at hour t , for principal component n (MW).

SOC	State Of Charge (%).
u_s^D	Binary variable to identify the operating segment s .
w	Variable used to linearize bi-linear terms.
$x^{C/D/ID}$	Binary variable indicating Charging/Discharging/Idle mode.
$x p_{n,t}^{X,Y}$	Binary variable to calculate the modulus of $p_{n,t}^{X,Y}$.
$\hat{\cdot}$	Affine representation of an uncertain variable.
Functions	
$AFR^C(SOC)$	Charging air flow rate as a function of the SOC ($kg.MW^{-1}.s^{-1}$).
$AFR^D(P^D)$	Discharging air flow rate as a function of P^D ($kg.MW^{-1}.s^{-1}$).
$f(P^X)$	Revenue equation as a function of the power dispatch (\$/h).
$HR(P^D)$	Heat rate as a function of P^D (GJ/MWh).
$\mathcal{F}$	Objective function.

I. INTRODUCTION

MAINTEINING demand-supply balance while operating securely is the main objective of power system operations [1]. With governments focusing on carbon emission reductions, there is a growing penetration of large-scale Renewable Energy Sources (RES) [2], but despite being environmentally friendly, RES are highly intermittent and uncertain, thus representing a volatile and difficult to forecast power source, which increases the complexity of operating power grids reliably. For this purpose, Energy Storage Systems (ESS) are being integrated into power systems to provide the necessary balancing and flexibility in system operation, storing energy to provide a variety of system services [3].

ESS can provide various services such as energy arbitrage, peak shaving, frequency regulation, etc., and with electricity prices varying over the day, with significantly different off-peak and peak prices, opportunities for profitable investments in ESS abound. Hence, ESS facilities can be a system asset or can be owned by private investors, and depending on the facility's capacity compared to the overall system, it may operate either as a price-taker or price-maker. For example, in [4], the operation and profit of a Battery Energy Storage System (BESS) is evaluated, participating as a price-taker in energy transactions and as a price-maker in ancillary services, considering uncertainties of energy prices and bids and offers from other participants. An assessment of non-strategic (price-taker) and strategic (price-maker) participation of a Pumped Hydro Storage (PHS) in energy and regulation markets is presented in [5], considering uncertainties on demand and competitors' offers, and assuming that the PHS can charge and discharge simultaneously.

Compressed Air Energy Storage (CAES) is a mechanical-based technology appropriate for bulk power systems applications. There are two diabatic facilities operating worldwide, the 290 MW plant built in Huntorf, Germany, since 1978 [6], and the 110 MW plant in McIntosh, Alabama, USA, since 1991 [7]. More recently, new CAES facilities are being deployed,

as for example the 1.75 MW adiabatic CAES in Goderich, Ontario, Canada, which can charge and discharge at the same time [8]. The bulk storage capacity and high efficiency with a low per-unit capacity cost of these CAES facilities is promising [9]. Thus, there have been recent CAES studies such as the mathematical model to represent CAES systems in steady-state and dynamic studies of power systems presented in [10], where the facility is considered to charge and discharge simultaneously, contributing to frequency regulation of the system. A self-scheduling model to maximize the daily profit of a price-taker CAES facility is presented in [11]; however, the thermodynamic characteristics of the facility are modeled as step functions, which increases the model complexity, and uncertainties are not considered. In [12], the authors proposed a self-scheduling model for the CAES system, where its thermodynamic characteristics are modeled as linear functions and uncertainties are considered using Robust Optimization (RO) and Affine Arithmetic (AA) techniques, thus yielding optimum schedules which are protected against uncertainties, while the computational burden is significantly lower with respect to [11].

To represent uncertainties, classical optimization methods have used probability density functions that, despite yielding a robust model, presents non-linear characteristics; besides, a large amount of data is required to achieve an accurate representation [13]. Thus, methods based on range-arithmetic, such as AA and RO, have been proposed in the literature. For example, in [14], AA is used to represent uncertainties in RES generation to solve the Optimum Power Flow (OPF) problem; the method is benchmarked with Monte Carlo Simulations (MCS), obtaining similar results with significantly lower computational burden. A microgrid Unit Commitment (UC) problem is modeled using AA to represent the uncertainties in demand and RES generation in [15], providing robust and cost-effective solutions with a reasonable computational burden. While AA is focused on optimizing the center and radius of the objective function, RO focuses on optimizing for the worst case scenario. However, classical RO methods provide non-adjustable solutions. Thus, some works propose the use of Affine Policies (AP) to improve RO results. For example, in [16], an adjustable robust OPF considering uncertainties in RES generation is proposed, which yields an optimum schedule, where AP adjusts the automatic generation control dispatch for a given mismatch in RES. In [17], an adaptive RO for UC of large systems, considering load and RES generation uncertainty, uses a simplified AP to obtain an adjustable schedule for system generators.

Power systems operation is sensitive to many parameters, which may lead to complex numerical analyses with high computational burden. Principal Components Analysis (PCA) has been employed by researchers to reduce the dimension of the problem. In [18], an AA-based framework for power flow and OPF studies is presented considering uncertainties in RES generation. Due to the large number of noise variables, PCA is employed to determine their main components, reducing the computation burden by approximately 80%, with an error tolerance of 10^{-3} . A multistage robust UC considering dynamic uncertainty sets and ESS is presented in [19] for a

large power system, using PCA to reduce the dimensionality of the uncertainty set, hence reducing the computational burden.

Although there are several studies implementing the aforementioned techniques for power systems studies, most are focused on operation of the power system such as power flow, OPF and UC [14]–[17], and only a few are focused on the facility owner's point of view [11], [12]. While one of the main contributions of [12] is a self-scheduling model for a price-taker CAES facility using AA to represent uncertainties, the results obtained, which are the center and radius of the profit and power dispatch, may become too conservative and yield a high radius, leading to low profit or even losses, in the most pessimistic scenario. To the best of the authors' knowledge, no works have been reported on applying both PCA and AP to a self-scheduling model of an ESS facility. Furthermore, since historical electricity price data presents several outliers such as price spikes and negative prices, yielding a wide interval of uncertainty, the representation of uncertainties using range arithmetic techniques is an issue. In this context, the main objectives and contributions of this paper are the following:

- Propose a PCA-AP-based model considering uncertainties in electricity prices and representing CAES facilities that can charge and discharge concurrently, considering the previously developed self-scheduling model for a CAES facility in [12]. This approach yields more accurate center and radius of the profit and power dispatch of the CAES facility, as the affine terms model the price uncertainties in more detail, considering inter-hour dependencies for each uncertain price, plus these terms are treated as unrestricted variables as opposed to the strictly positive numbers required in the AA formulation presented in [12], which results in tighter affine intervals.
- Apply the Box Cox transformation method to estimate the price intervals based on the historical electricity price data. This novel approach does not rely on assuming fixed intervals for price uncertainties for each time interval, unlike [12], and is thus more accurate and practical.
- Benchmark the results obtained using the proposed PCA-AP-based model with respect to the results from existing AA and MCS-based approaches. This allows for a comprehensive analysis of uncertainty-based approaches, demonstrating that the proposed PCA-AP technique yields tighter profit intervals compared with the AA technique proposed in [12], thus producing better margins of operation for a given range of uncertainty.

The rest of the paper is organized as follows: Section II presents a brief background to the CAES deterministic model considering the thermodynamic characteristics, and the AA formulation for the CAES model. The PCA formulation and PCA-AP-based model for the CAES facility are presented in Section III. Results and comparisons obtained from the case study are presented in Section IV, and Section V highlights the main conclusions and contributions of the paper.

II. BACKGROUND

A. CAES Deterministic Model

Adopting the model proposed in [12], with some relevant modifications highlighted throughout this section, the CAES facility is assumed to be a price-taker, i.e., it does not affect electricity market prices by its actions, thus representing existing and under-development CAES facilities, which makes the proposed approach timely and relevant in practice. Participating in the day-ahead market, providing energy, spinning and idle reserves, the objective is to maximize the daily profit of the facility, given as follows:

$$\max_{P_t^X} \mathcal{F} = \sum_t^T [f(P_t^X) - OC_t] \quad (1)$$

where:

$$f(P_t^X) = A_t \pi_t^E + B_t \pi_t^{SR} + P_t^{ID} \pi_t^{ID} \quad \forall t \quad (2)$$

$$A_t = P_t^D - P_t^C \quad \forall t \quad (3)$$

$$B_t = P_t^{SRD} + P_t^{SRC} \quad \forall t \quad (4)$$

$$OC_t = CO_t^{NG} + P_t^D VOM^e + P_t^C VOM^c \quad \forall t \quad (5)$$

$$CO_t^{NG} = HR(P_t^D) P_t^D \pi^{NG} \quad \forall t \quad (6)$$

$$P_t^X \geq 0 \quad \forall t, X \quad (7)$$

All variables and parameters are defined in the Nomenclature section. The revenue from participation in the energy market is given in (2), where the CAES facility pays for charging and receives a payment for discharging (3), spinning reserve (4), and idle reserve services. The operations cost is given in (5), where the first term denotes the cost of natural gas usage in the discharging stage (6), and the second and third terms denote the Variable Operational and Maintenance (VOM) costs for the expander and compressor, respectively.

The operational constraints can be defined as follows:

$$x_t^C + x_t^D \leq 1 \quad \forall t \quad (8)$$

$$x_t^C + x_t^{ID} \leq 1 \quad \forall t \quad (9)$$

$$x_t^D + x_t^{ID} \leq 1 \quad \forall t \quad (10)$$

$$P_t^C \leq \bar{P}^C x_t^C \quad \forall t \quad (11)$$

$$P_t^{SRC} \leq P_t^C - \underline{P}^C x_t^C \quad \forall t \quad (12)$$

$$P_t^{SRD} \leq \bar{P}^D x_t^D - P_t^D \quad \forall t \quad (13)$$

$$P_t^D \geq \underline{P}^D x_t^D \quad \forall t \quad (14)$$

$$0 \leq P_t^{ID} \leq QSC x_t^{ID} \quad \forall t \quad (15)$$

In (8), (9), and (10), the CAES is set to operate in either charging, discharging, or idle modes; removing constraint (8) would allow to represent CAES facilities that can charge and discharge at the same time, which is not possible in the model introduced in [12]. Equations (11) to (15) represent the minimum and maximum power capacities of charging, discharging, spinning, and idle reserves, respectively. In (11) and (12), the lower and upper limits of the compressor charging power should be between 40%-100% of the rated power, where its efficiency is approximately constant, whereas for (13) and (14) the lower and upper limits of the turbine discharging power

should be between 30%-100% of its rated power, where its efficiency is not further compromised [20].

The inter-temporal variation of the State of Charge (SOC) can be represented by:

$$SOC_{t+1} = SOC_t + \frac{air_t^C 3,600}{CA} - \frac{air_t^D 3,600}{CA} \quad \forall t \quad (16)$$

$$air_t^C = AFR^C(SOC_t)P_t^C \quad \forall t \quad (17)$$

$$air_t^D = AFR^D(P_t^D)P_t^D \quad \forall t \quad (18)$$

$$SOC_{t+1} \geq SOC^f \quad \forall t = T \quad (19)$$

The SOC at time $t+1$ is calculated in (16). The amount of air charged/discharged is determined by the Air Flow Rate AFR of the facility multiplied by the charging/discharging power, as presented in (17) and (18), respectively. Since the facility participates in the day-ahead market, it requires a minimum SOC in order to be better prepared, which is represented in (19) by SOC^f .

The thermodynamic model discussed next is based on the Huntorf project [6], with Fig. 1 presenting the thermodynamics characteristics of the CAES facility. As the SOC of the facility increases, so does the pressure inside the cavern; hence, it becomes difficult to store more air inside and thus AFR^C decreases, as illustrated in Fig.1a. The efficiency of the high pressure turbine reduces when the facility operates below its rated power; thus, a greater AFR^D is required to generate one unit of power, as illustrated in Fig.1b, where the rated power is considered to be 100 MW. A higher AFR requires a greater amount of natural gas, thus, the Heat Rate HR of the facility increases for lower discharge, as presented in Fig.1c.

Since (17) is non-linear, McCormick Envelopes can be used to linearize it in a slightly relaxed form. Thus, based in [12] and Fig. 1, the mass of air charged at time t can be represented as follows:

$$air_t^C = -0.3w_t + 1.95P_t^C \quad \forall t \quad (20)$$

$$w_t \leq Mx_t^C \quad \forall t \quad (21)$$

$$w_t \geq \underline{SOC}P_t^C + SOC_tP^C - \underline{SOC}P^C - M(1 - x_t^C) \quad \forall t \quad (22)$$

$$w_t \geq \overline{SOC}P_t^C + SOC_tP^C - \overline{SOC}P^C - M(1 - x_t^C) \quad \forall t \quad (23)$$

$$w_t \leq \underline{SOC}P_t^C + SOC_tP^C - \underline{SOC}P^C + M(1 - x_t^C) \quad \forall t \quad (24)$$

$$w_t \leq \overline{SOC}P_t^C + SOC_tP^C - \overline{SOC}P^C + M(1 - x_t^C) \quad \forall t \quad (25)$$

Furthermore, the air discharged in (18) and the cost of natural gas in (6) can be represented from Fig. 1 as follows:

$$air_t^D = -\frac{0.9}{70}P_t^{D^2} + \frac{18.8}{7}P_t^D \quad \forall t \quad (26)$$

$$\frac{CO_t^{NG}}{\pi^{NG}} = -\frac{1.25}{70}P_t^{D^2} + \frac{457.5}{70}P_t^D \quad \forall t \quad (27)$$

Taking the derivative of each equation in terms of P_t^D and equating to zero yields the maxima, which are 104.4 MW and 183 MW, respectively. Since the CAES facility operates

between 30-100 MW, the air discharged can be represented by (26), with the cost of natural gas (27) being in its linear region; thus, these equations can be approximated using piecewise linearization and a straight line, respectively, as follows (Fig. 2):

$$P_t^D = \sum_{s=1}^2 (b_{t,s}^D + \underline{q}_s u_{t,s}^D) \quad \forall t \quad (28)$$

$$\sum_{s=1}^2 u_{t,s}^D = x_t^D \quad \forall t \quad (29)$$

$$0 \leq b_{t,s}^D \leq \bar{b}^D u_{t,s}^D \quad \forall t, s \quad (30)$$

$$air_t^D = \sum_{s=1}^2 (ang_s^D b_{t,s}^D + \underline{D}_s u_{t,s}^D) \quad \forall t \quad (31)$$

$$CO_t^{NG} = (53.57x_t^D + 4.215P_t^D)\pi^{NG} \quad \forall t \quad (32)$$

In (28) to (30), the power discharged by the CAES facility is decomposed in terms of variables and parameters that identify the point of operation of the facility, which is used to determine the mass of air discharged (31). In (32), the cost of natural gas can be determined from the state of operation and discharged power, which simplifies and reduces computational burden with respect to the representation of these costs in [12].

B. AA Formulation of CAES Operation Model

1) *AA Overview:* As proposed in [21], AA is a range arithmetic technique that keeps track of correlated uncertainties between variables yielding a solution with real-time adjustment. In this case, an uncertain variable χ is represented in its affine form as follows:

$$\hat{\chi} = \chi_0 + \chi_1\varepsilon_1 + \chi_2\varepsilon_2 + \cdots + \chi_p\varepsilon_p = \chi_0 + \sum_{h=1}^p \chi_h\varepsilon_h \quad (33)$$

where χ_0 denotes the center and most likely value of the variable χ ; ε_h are the noise parameters in the interval $[-1, 1]$; and χ_h are the magnitudes of the corresponding uncertainty components.

When using AA, there are two sources of uncertainties: external and internal. Imprecise data and uncertainty in the mathematical model are external sources, while truncation and round-off errors are internal sources of uncertainty. Calculations involving AA can be affine or non-affine operations; thus, considering two affine quantities, $\hat{a} = a_0 + \sum_{h=1}^p a_h\varepsilon_h$ and $\hat{b} = b_0 + \sum_{h=1}^p b_h\varepsilon_h$, and a parameter α , the affine operations can be defined as follows:

$$\hat{z} = \hat{a} \pm \hat{b} = (a_0 \pm b_0) + \sum_{h=1}^p (a_h \pm b_h)\varepsilon_h \quad (34)$$

$$\hat{z} = \alpha\hat{a} = \alpha a_0 + \alpha \sum_{h=1}^p a_h\varepsilon_h \quad (35)$$

$$\hat{z} = \hat{a} \pm \alpha = (a_0 \pm \alpha) + \sum_{h=1}^p a_h\varepsilon_h \quad (36)$$

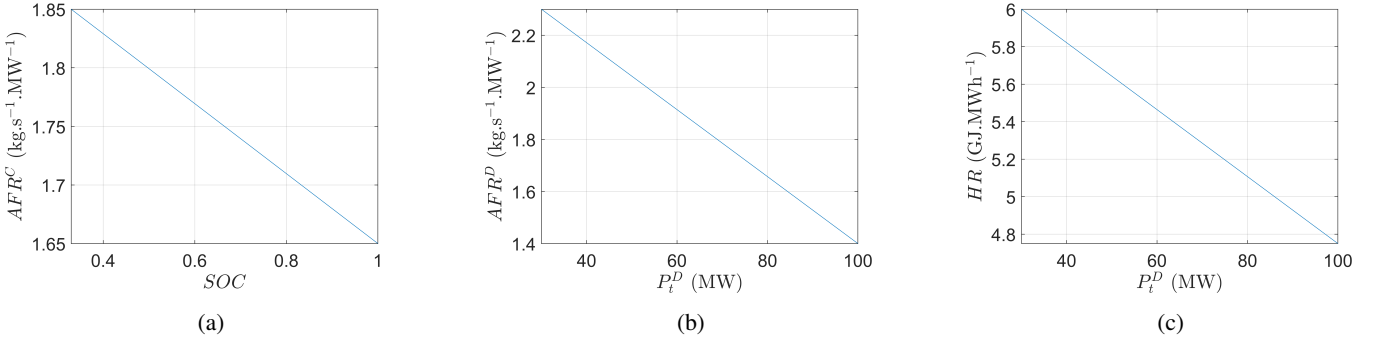


Fig. 1: CAES thermodynamic characteristics [12].

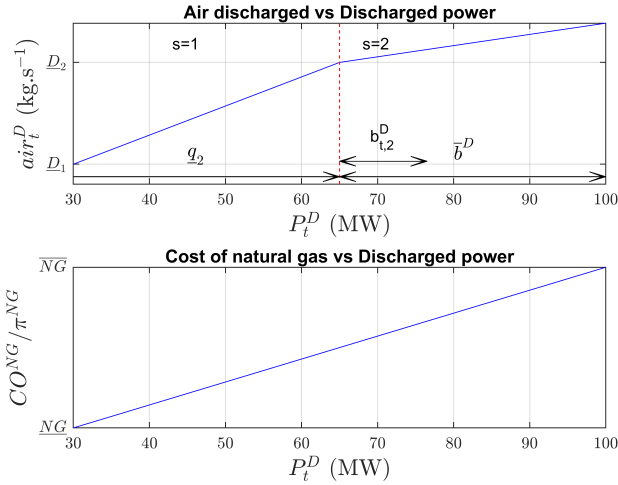


Fig. 2: Linear approximation of the quadratic functions.

and a non-affine operation is:

$$\hat{z} = \hat{a}\hat{b} = a_0b_0 + \sum_{h=1}^p (a_0b_h + b_0a_h)\varepsilon_h + z_k\varepsilon_k \quad (37)$$

where the resulting variable $\hat{z}$ contains information from $\hat{a}$ and $\hat{b}$, and the approximation error $z_k\varepsilon_k$.

The general representation for AA in an optimization problem is given by [18]:

$$\max_{\hat{z}} \hat{\mathcal{F}}(\hat{z}) \quad (38)$$

$$\text{s.t. } \hat{g}_l(\hat{z}) \stackrel{A}{\approx} 0 \quad \forall l \in L \quad (39)$$

$$\hat{h}_m(\hat{z}) \stackrel{A}{\leq} 0 \quad \forall m \in M \quad (40)$$

where the equality $\stackrel{A}{\approx}$ and inequality $\stackrel{A}{\leq}$ constraints are defined in the AA domain for the aforementioned AA quantities $\hat{a}$ and $\hat{b}$ as follows:

$$\stackrel{A}{\approx}: a_0 = b_0 \wedge a_h = b_h \quad \forall h \quad (41)$$

$$\stackrel{A}{\leq}: a_0 + \sum_{h=1}^p |a_h| \leq b_0 - \sum_{h=1}^p |b_h| \quad (42)$$

Thus, (41) indicates that one affine variable is equal to the other if all their terms are equal, and (42) states that one affine

variable is less than the other if the upper boundary of the left-hand side is lower than the lower boundary of the right-hand side.

2) *AA Model of the CAES Facility*: As presented in [12] for the self-scheduling model of the CAES facility, the uncertain market prices can be represented as follows:

$$\hat{\pi}_t^Y = \pi_{0,t}^Y (1 + \varepsilon_t^Y \overline{\Delta\pi_t^Y}) \quad \forall t, Y \quad (43)$$

where $\pi_{0,t}^Y$ denotes the price forecast. On the other hand, the power dispatch variables can be defined as follows:

$$\hat{P}_t^X = P_{0,t}^X + P_{1,t}^X \varepsilon_t^E + P_{2,t}^X \varepsilon_t^{SR} + P_{3,t}^X \varepsilon_t^{ID} \quad \forall t, X \quad (44)$$

$$P_{0,t}^X, P_{h,t}^X \geq 0 \quad \forall t, X, h \quad (45)$$

where ε_t^Y are predefined intervals $[-1, 1] \forall t$ and are thus known for all time periods considered, and the variables $P_{h,t}^X \forall h, t$ are computed using the AA model, explained next, so that $\hat{P}_t^X$ conservatively captures all possible facility dispatches accounting for the price uncertainties defined in (43) for each time interval considered, while accounting for all the constraints of the CAES facility, in order to avoid infeasibility in the solution.

The product of electricity prices with power dispatch in their affine forms, i.e., (43) and (44) yields the following:

$$\begin{aligned} \hat{P}_t^X \hat{\pi}_t^Y &= (P_{0,t}^X + \sum_{h=1}^p P_{h,t}^X \varepsilon_t^h) \pi_{0,t}^Y (1 + \varepsilon_t^Y \overline{\Delta\pi_t^Y}) \\ \hat{P}_t^X \hat{\pi}_t^Y &= \underbrace{P_{0,t}^X \pi_{0,t}^Y}_{\mathcal{F}_0(\hat{P}_t^X)} + \underbrace{P_{0,t}^X \pi_{0,t}^Y \overline{\Delta\pi_t^Y} \varepsilon_t^Y + \sum_{h=1}^p P_{h,t}^X \pi_{0,t}^Y \varepsilon_t^h}_{\mathcal{F}_A(\hat{P}_t^X)} \\ &\quad + \underbrace{\sum_{h=1}^p P_{h,t}^X \pi_{0,t}^Y \overline{\Delta\pi_t^Y} \varepsilon_t^h \varepsilon_t^Y}_{\mathcal{F}_{NA}(\hat{P}_t^X)} \quad \forall t, X, Y \end{aligned} \quad (46)$$

Substituting (43) and (44) in (1), executing the aforementioned affine and non-affine operations, and splitting the terms as per (46) the objective function can be formulated as follows:

$$\max_{\hat{P}_t^X} \sum_{t=1}^T \left(\mathcal{F}_0(\hat{P}_t^X) + \left| \mathcal{F}_A(\hat{P}_t^X) \right| - \left| \mathcal{F}_{NA}(\hat{P}_t^X) \right| \right) \quad (47)$$

where the first $\mathcal{F}_0$, second $\mathcal{F}_A$, and third terms $\mathcal{F}_{NA}$ denote the center, affine, and non-affine terms, respectively. Note that

the center and radius, which comprises $\mathcal{F}_A$ and $\mathcal{F}_{NA}$, are maximized for the center and affine terms, since the latter represent the external source of uncertainties, and minimize the non-affine terms, which represent the internal source of uncertainties, i.e., the product of two affine terms. This yields a single and reasonably conservative interval for the daily schedule of the CAES facility. Observe that a more conservative approach would be to increase the radius by maximizing the affine and non-affine terms, however, a large radius is undesirable since in pessimistic scenarios, it may lead to significantly lower profits or even losses. Thus, as discussed in [12], the model uses a risk-taking strategy by maximizing the profit center and the modulus of the radius of the affine terms and avoid feasibility issues associated with integer variables, as mentioned in [15], while minimizing the radius of non-affine terms to reduce the size of the interval to avoid being too conservative. Note that like any dispatch model, there is no guarantee of feasibility; nevertheless, with the proposed single-stage optimization approach, there have not been infeasibility issues encountered in the multiple simulations carried out. The main constraints of the AA mathematical model are presented in Appendix A

III. PCA-AP-BASED CAES MODEL

The model proposed here is based on a PCA approach, to reduce the complexity of the mathematical model. Thus, PCA is used to reduce the dimension of the large data sets of prices and thus discover potential relationships among variables. This is accomplished as follows, as per [22]:

- **Standardization:** To avoid dominance of variables with greater magnitudes, the original data d is standardized using the $\mathcal{Z}$ -score [23]:

$$\mathcal{Z} = \frac{1}{\sigma}(d - d_{av}\mathbf{1}) \quad (48)$$

where d_{av} is the average and σ the standard deviation of d , and $\mathbf{1}$ a vector of ones.

- **Covariance:** Compute the covariance matrix to determine how much the variables vary with respect to each other, thus capturing their inter-relationships.
- **Eigen-analysis:** The eigenvectors $\mathbf{V}$ of the covariance matrix are the Principal Components (PC) direction and the corresponding eigenvalues $\mathbf{\Lambda}$ yield their variance, as follows:

$$\mathbf{Z} = \text{cov}(\mathcal{Z}) \quad (49)$$

$$[\mathbf{V}, \mathbf{\Lambda}] = \text{eig}(\mathbf{Z}) \quad (50)$$

$$\lambda = \text{diag}(\mathbf{\Lambda}) \quad (51)$$

$$\lambda_{inf} = \frac{1}{\sum_i \lambda_i} \lambda \quad (52)$$

where λ_{inf} represents the information level, in terms of percentage, for each eigenvector.

- **Feature Vector:** Discarding the PCs with lesser information from $\mathbf{V}$, the remaining ones form a matrix known as Feature Vector $\mathbf{F}$, with the PCs being calculated as follows:

$$PC = \mathbf{F}^T d \quad (53)$$

where the quantity of PCs is determined by sorting λ_{inf} in descending order and aggregating them until a minimum amount of information is achieved.

Therefore, the first step to obtain the PCA-AP model of the CAES facility is to obtain $\mathbf{V}$. Hence, from the hourly electricity prices for D days, the price matrix $\mathbf{\Pi}^Y$ can be expressed as:

$$\mathbf{\Pi}^Y = \begin{bmatrix} \pi_{1,1}^Y & \cdots & \pi_{1,24}^Y \\ \vdots & \ddots & \vdots \\ \pi_{D,1}^Y & \cdots & \pi_{D,24}^Y \end{bmatrix}_{D \times 24} \quad \forall Y \quad (54)$$

where the rows represent the day and the columns the hour. Executing the aforementioned PCA steps, the covariance matrix $\mathbf{Z}_{24 \times 24}$ in (49) can be determined with its associated eigenvectors $\mathbf{V}^Y$ and eigenvalues $\mathbf{\Lambda}^Y$, and the PCs can then be determined as follows:

$$PC_{24 \times 1}^Y = (\mathbf{V}^Y)^T \pi_{24 \times 24}^Y \quad \forall Y \quad (55)$$

Electricity prices have a forecast value and a range of uncertainty, which can be represented using AP as follows:

$$\hat{\pi}_t^Y = \pi_{0,t}^Y (1 + \Delta\pi_t^Y) \quad \forall t, Y \quad (56)$$

$$-\overline{\Delta\pi}_t^Y \leq \Delta\pi_t^Y \leq \overline{\Delta\pi}_t^Y \quad \forall t, Y \quad (57)$$

Note that (56) is different from the AA affine model (43), since the price variation $\Delta\pi_t^Y$ is now a variable instead of a parameter $\overline{\Delta\pi}_t^Y$. Replacing (56) in (55), the AP form of PC^Y can be obtained:

$$\widehat{PC}^Y = \underbrace{(\mathbf{V}^Y)^T \pi_0^Y}_{PC_0^Y} + \underbrace{(\mathbf{V}^Y)^T \pi_0^Y \odot \Delta\pi^Y}_{\Delta PC^Y} \quad \forall Y \quad (58)$$

where $\pi_0^Y = [\pi_{0,1}^Y \pi_{0,2}^Y \cdots \pi_{0,24}^Y]^T$ and $\pi_0^Y \odot \Delta\pi^Y = [\pi_{0,1}^Y \Delta\pi_1^Y \pi_{0,2}^Y \Delta\pi_2^Y \cdots \pi_{0,24}^Y \Delta\pi_{24}^Y]^T$. From (58), the following limits can be defined, based on (57):

$$\overline{\Delta PC}^Y = (\mathbf{V}^Y)^T \pi_0^Y \odot \overline{\Delta\pi}^Y \quad \forall Y \quad (59)$$

$$-\overline{\Delta PC}_n^Y \leq \Delta PC_n^Y \leq \overline{\Delta PC}_n^Y \quad \forall n, Y \quad (60)$$

Hence, the hourly electricity prices can be represented as follows:

$$\mathbf{L}^Y = \begin{bmatrix} l_{1,1}^Y & \cdots & l_{1,24}^Y \\ \vdots & \ddots & \vdots \\ l_{24,1}^Y & \cdots & l_{24,24}^Y \end{bmatrix} = ((\mathbf{V}^Y)^T)^{-1} \quad \forall Y \quad (61)$$

Discarding the columns associated with the PCs with low information, the AP form of prices can be represented as a function of PCs as follows:

$$\hat{\pi}_{24 \times 1}^Y = \mathbf{L}_{24 \times NPC}^Y \widehat{PC}_{NPC \times 1}^Y \quad \forall Y \quad (62)$$

$$\hat{\pi}_t^Y = \underbrace{\sum_{n=1}^{NPC} l_{t,n}^Y PC_{0,n}^Y}_{\pi_{0,t}^Y} + \underbrace{\sum_{n=1}^{NPC} l_{t,n}^Y \Delta PC_n^Y}_{\Delta\pi_t^Y} \quad \forall t, Y \quad (63)$$

And the power variables in their AP form can be represented as follows:

$$\hat{P}_t^X = P_{0,t}^X + \sum_Y \sum_{n=1}^{NPC} p_{n,t}^{X,Y} \Delta PC_n^Y \quad \forall t, X \quad (64)$$

where $p_{n,t}^{X,Y}$ are unrestricted-sign variables, and:

$$P_{0,t}^X + \sum_Y \sum_{n=1}^{NPC} |p_{n,t}^{X,Y}| \overline{\Delta PC_n^Y} \leq \bar{P}^X \quad \forall t, X \quad (65)$$

$$P_{0,t}^X - \sum_Y \sum_{n=1}^{NPC} |p_{n,t}^{X,Y}| \overline{\Delta PC_n^Y} \geq \underline{P}^X \quad \forall t, X \quad (66)$$

Since there are modulus in the formulation, to avoid non-linearities, auxiliary variables and binary variables have to be introduced by applying the “big M” linearization technique [24], obtaining an equivalent formulation for the problem, as follows:

$$p_{n,t}^{X,Y} = p_{n,t}^{X,Y+} - p_{n,t}^{X,Y-} \quad \forall t, n, X, Y \quad (67)$$

$$p_{n,t}^{X,Y+} \leq M x p_{n,t}^{X,Y} \quad \forall t, n, X, Y \quad (68)$$

$$p_{n,t}^{X,Y-} \leq M(1 - x p_{n,t}^{X,Y}) \quad \forall t, n, X, Y \quad (69)$$

$$|p_{n,t}^{X,Y}| = p_{n,t}^{X,Y+} + p_{n,t}^{X,Y-} \quad \forall t, n, X, Y \quad (70)$$

$$p_{n,t}^{X,Y+}, p_{n,t}^{X,Y-} \geq 0 \quad \forall t, n, X, Y \quad (71)$$

Hence, as the number of principal components NPC increase, so does number of variables and constraints in the model.

The product of (63) with (64) yields the following terms:

$$\begin{aligned} \hat{P}_t^X \hat{\pi}_t^Y &= \left(P_{0,t}^X + \sum_Y \sum_{n=1}^{NPC} p_{n,t}^{X,Y} \Delta PC_n^Y \right) (\pi_{0,t}^Y + \Delta \pi_t^Y) \\ &= \underbrace{P_{0,t}^X \pi_{0,t}^Y}_{\mathcal{F}_{0,t}} + \underbrace{P_{0,t}^X \Delta \pi_t^Y + \sum_Y \sum_{n=1}^{NPC} p_{n,t}^{X,Y} \Delta PC_n^Y \pi_{0,t}^Y}_{\mathcal{F}_{l,t}} \\ &\quad + \underbrace{\sum_Y \sum_{n=1}^{NPC} p_{n,t}^{X,Y} \Delta PC_n^Y \Delta \pi_t^Y}_{\mathcal{F}_{nl,t}} \quad \forall t, X, Y \quad (72) \end{aligned}$$

Substituting (63) and (64) in (1), executing the mathematical operations, and splitting the terms as in (72) the objective function can be represented as follows:

$$\max_{\hat{P}_t^X} \sum_{t=1}^T \left(\mathcal{F}_{0,t} + \mathcal{F}_{l,t} - \mathcal{F}_{nl,t} \right) \quad (73)$$

where the resulting terms are divided into center $\mathcal{F}_{0,t}$, linear $\mathcal{F}_{l,t}$, and non-linear $\mathcal{F}_{nl,t}$ terms; the latter two terms comprise the radius of profit. The model is optimized for the most conservative scenario, i.e., $\Delta PC_n^Y = \overline{\Delta PC_n^Y}$, to linearize (72). However, since the objective is to maximize the daily profit of the CAES facility, to avoid a conservative solution, the model is maximized for the center and linear terms and minimized for the non-linear terms, to avoid a large radius that may yield economic losses in pessimistic scenarios. Constraints (2)-(5), (8)-(16), (21)-(25), (28)-(32), (63)-(64), together with (73) fully define the proposed PCA-AP-based Mixed Integer Linear Programming model of the CAES facility. The main constraints of the PCA-AP mathematical model are presented in Appendix B.

IV. RESULTS AND DISCUSSIONS

The simulation studies presented next are based on the Huntorf facility [6], using the thermodynamic model explained in Section II-A and based on Figs. 1 and 2, which is the same CAES system used in [12], using the same price data to allow direct comparisons. Thus, the maximum charging and discharging capacities of the CAES facility are considered to be 60 MW and 100 MW, respectively, while the minimum charging and discharging capacities are 25 MW and 30 MW, respectively, with a quick-start capacity of 40 MW. The SOC is maintained between $\underline{SOC} = 33\%$ and $\overline{SOC} = 100\%$, the initial SOC is $SOC_1 = 60\%$ and the final SOC is set to be at least $SOC^f = 60\%$. The maximum mass of air the cavern can store is assumed to be the amount of air discharged by the plant operating for eight continuous hours at full power. The energy, spinning and idle reserve prices used in the studies are obtained from electricity price data for 2015 to 2019 [25], with all prices adjusted to 2019 using the average inflation index of Canada for the 5 years taken [26]. The simulations are carried out for the months of February, May, August, and November, which are relevant months for each season. The price intervals for the simulations are obtained as follows:

- Average daily electricity prices are transformed using the Box Cox Transformation, considering that normality is an important assumption for statistical techniques. However, real world data may not always adhere to a normal shape behavior; thus, the Box Cox transformation can be employed to transform the non-normal data into a normal shape [27]. The Box Cox transformation W of non-normal data d is determined as follows:

$$W = \begin{cases} \frac{d^\gamma - 1}{\gamma}, & \text{if } \gamma \neq 0 \\ \log(d), & \text{if } \gamma = 0 \end{cases} \quad \gamma \in [-5, 5] \quad (74)$$

where γ is a parameter estimated to minimize the standard deviation of the transformed data. Given the transformation W , the extreme points are filtered from the normal distribution using the concept of Confidence Interval (CI). Sorting the values of W in ascending order W^* , and assuming that the total number of samples is S , a CI of $P\%$ is given by [28]:

$$CI_{P\%} = \left[\bar{W} - (W_{(1+P)/2\%}^* - \bar{W}), \bar{W} - (W_{(1-P)/2\%}^* - \bar{W}) \right] \quad (75)$$

where $\bar{W}$ is the average of W ; $W_{(1-P)/2\%}^*$ is the $(1 - P)/2$ floor $W_{\lfloor S(1-P)/2 \rfloor}^*$ of the S samples; and $W_{(1+P)/2\%}^*$ is the $(1 + P)/2\%$ ceiling $W_{\lceil S(1+P)/2 \rceil}^*$ of the S samples. Hence, the remaining data represents the majority of occurrences in the original data, with a CI of $P = 50\%$ being used to filter the significant price spikes.

- From the remaining data, the average and standard deviation of the hourly prices can be determined, i.e., the center and interval of uncertainty, respectively.

To avoid a conservative interval of uncertainty, an upper limit on the price deviations is set to $\overline{\Delta \pi_t^Y} = 15\% \forall t, Y$ [29]. This procedure allows for a more accurate and realistic modeling of the price uncertainties. The filtered price intervals obtained

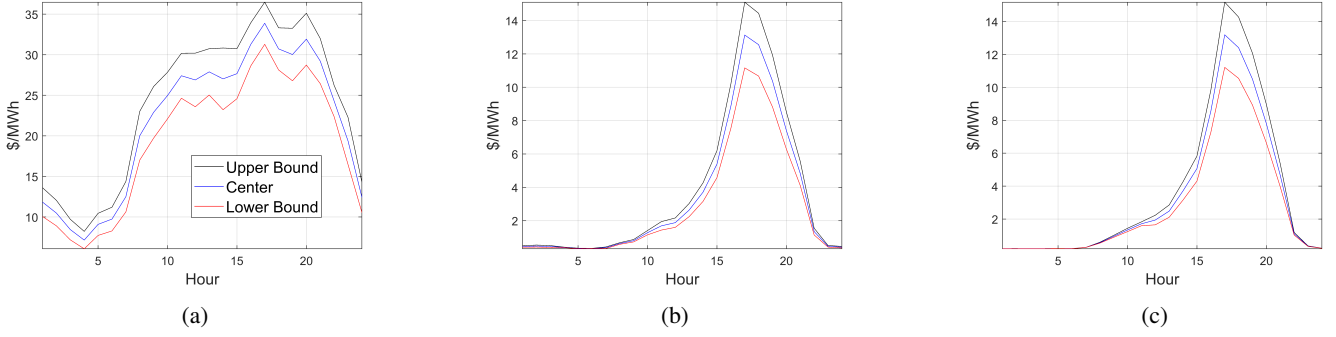


Fig. 3: Price intervals for the month of August: (a) Energy, (b) Spinning, and (c) Idle.

TABLE I: NPC and Electricity Prices RMSE for August

PCA	π^E		π^{SR}		π^{ID}	
	NPC	RMSE	NPC	RMSE	NPC	RMSE
95%	16	2.7458	15	0.5815	15	0.4929
97.5%	19	2.2941	18	0.2370	18	0.4508
99%	21	1.9980	20	0.1927	21	0.2217
100%	24	0	24	0	24	0

for the month of August, which is the peak demand month, are presented in Fig. 3. All simulations were performed in GAMS interfaced with MATLAB.

A. PCA-AP Results

The PCA-AP model simulations were carried out for different information levels, as per (52) and (53), of 95%, 97.5%, 99%, and 100%. Since electricity prices are represented as a function of PCs, discarding those that represent lesser information introduce an error associated with the PC representation and original data. Thus, for each information level, the NPC and the root mean square error (RMSE) for each electricity price, the center and radius of the profit, the computational burden, and optimum schedules are computed.

Table I presents the NPC and RMSE for the electricity prices of one of the four months considered in the studies; similar results were obtained for the other three months. Note that for higher information levels, the NPC increases and the RMSE decreases, as expected. A comparison between the energy price representation obtained from PCA with the original data for August is presented in Fig. 4. Note that for information levels of 95% and 97.5%, the energy prices obtained differ significantly from the original data, especially for the former.

Table II presents the center value and radius of the CAES facility's profit, number of constraints and variables, and computational burden for each month. Note that in all cases the center is greater than the radius, hence, there are no losses in the most pessimistic scenario. Since the 95% and 97.5% representations involve fewer variables and constraints as compared to the 99% and 100% cases, they would be expected to be computationally more efficient; however, that is not the case, as the CAES model contains several binary variables, and therefore it is sensitive to the input data. Hence, discarding the PCs with less information yielded a poor representation of

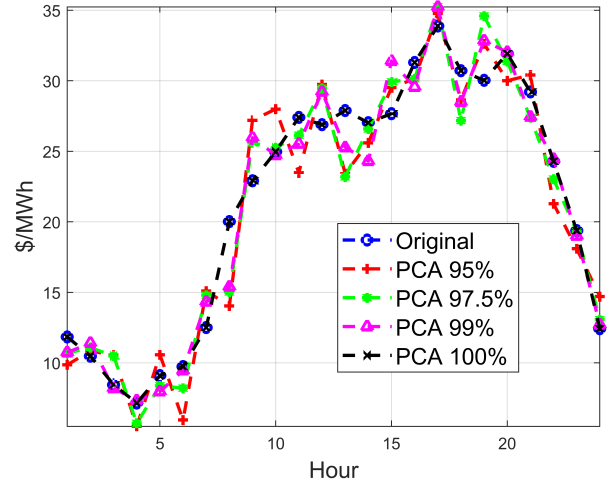


Fig. 4: August energy prices from PCA and original data.

the input data, which consequently affected the accuracy of the model and its computational costs, despite fewer variables and constraints, and thus it was not possible to obtain a consistent pattern of intervals and computational costs for all months. Note that, even though the computational burden varies depending on the monthly price profile, all simulations converged in less than 5 minutes, which is reasonable for day-ahead operation planning.

The optimum SOC profiles for the CAES facility for August, considering the different levels of information, are presented in Fig. 5. The depicted SOC results were obtained from the real-time price mismatch $\Delta\pi_t^Y$, with the deviation in the PCs ΔPC^Y being determined as per (58), and the dispatch being updated as per (64). Note that there are some differences in the hourly dispatch, however, they do not differ significantly from one another.

B. AA Results

The AA model yields a single optimum schedule given the price intervals discussed in Fig.3. The center and modulus of radius of the CAES facility's profit are determined as per (47). Thus, Table III presents the center and radius of the facility's profit and the computational burden for all AA simulations. Note that in all cases the center is greater than the radius.

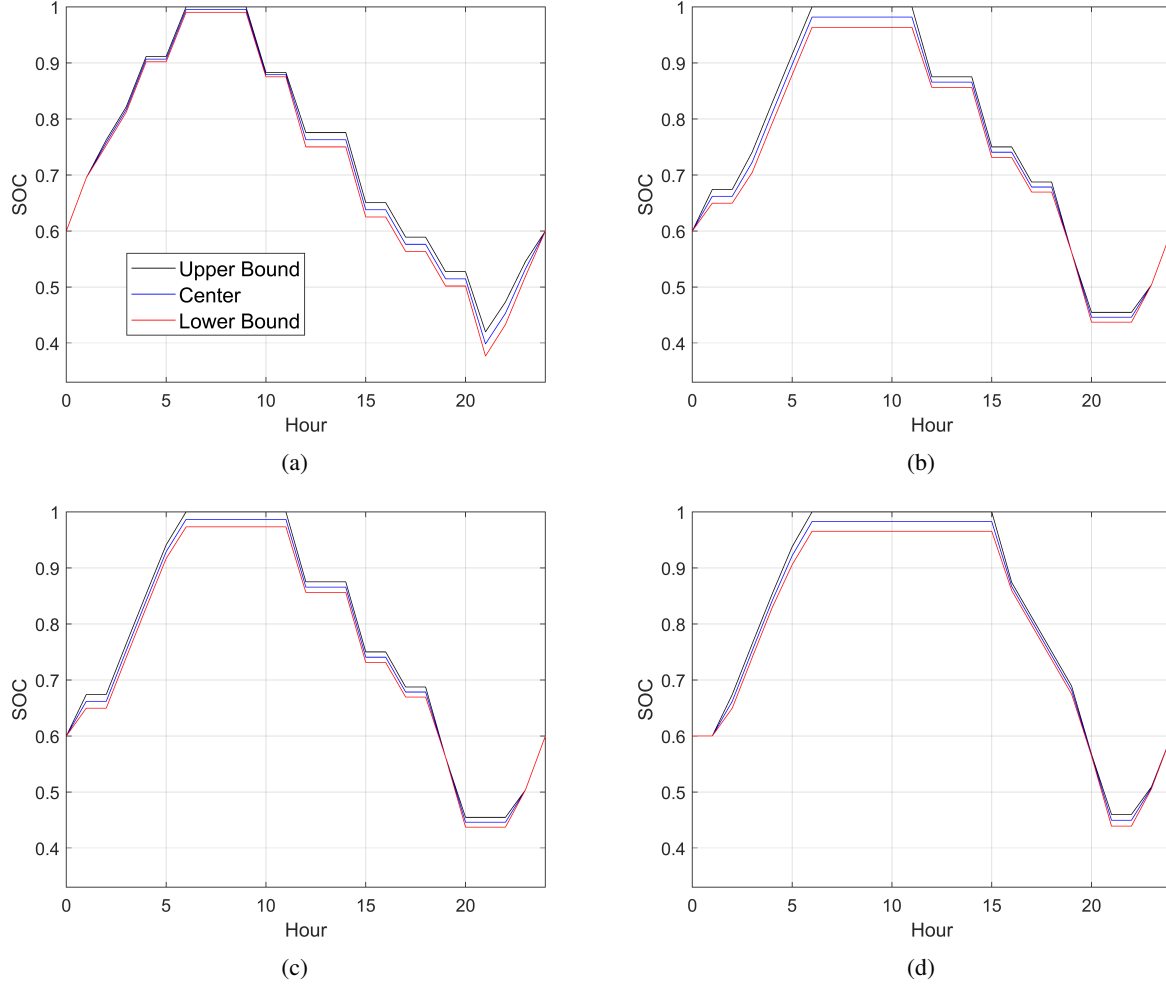


Fig. 5: SOC profiles from PCA-AP Schedules for August: (a) 95%, (b) 97.5%, (c) 99%, and (d) 100%.

TABLE II: PCA-AP Profit and Computational Burden

Month	PCA	Profit (\$)		Constraints	Variables	Burden
		Center	Radius			
Feb.	95%	7,603	4,651	36,609	36,656	1min11s
	97.5%	8,716	2,368	42,369	42,416	47s
	99%	7,848	2,695	46,689	46,736	56s
	100%	8,399	2,186	53,169	53,216	47s
May	95%	10,977	2,570	33,009	33,056	6.37s
	97.5%	10,583	2,494	39,489	39,536	7.47s
	99%	10,562	2,429	45,249	45,296	26s
	100%	10,505	2,419	44,359	8,857	41s
Aug.	95%	5,921	2,341	34,449	34,496	3min25s
	97.5%	5,977	2,341	40,929	40,976	2min38s
	99%	6,210	2,151	45,969	46,016	3min38s
	100%	5,637	2,127	53,169	53,216	4min17s
Nov.	95%	6,469	1,838	33,729	33,776	33s
	97.5%	6,374	2,299	39,489	39,536	29s
	99%	6,225	2,205	45,249	45,296	41s
	100%	6,383	1,984	53,169	53,216	46s

TABLE III: AA Profit and Computational Burden

Month	Profit (\$)		Burden
	Center	Radius	
February	7,322	5,351	12s
May	10,237	2,616	2.40s
August	5,695	3,190	48s
November	6,186	3,048	3.84s

Hence, in the most pessimistic scenario there are no losses. Observe that all the cases show a low computational burden, converging in less than a minute.

The optimum SOC profile for August obtained from the AA based charging/discharging dispatch of the CAES facility is illustrated in Fig. 6. These were obtained based on the real-

time price mismatch $\Delta\pi_t^Y$, from which the noise terms can be defined as $\varepsilon_t^Y = \frac{\Delta\pi_t^Y}{\Delta\pi_t}$, and the dispatch is then determined as per (44).

C. Comparisons

To validate the results obtained from the PCA-AP and AA models, both were benchmarked using the MCS approach. For this purpose, three random uniform distribution functions, with 3,000 data points each, were generated to yield different combinations of electricity prices as input data. The convergence tolerance for the MCS was set to 0.1%.

Table IV presents the lower boundary (LB) and upper boundary (UB) of the CAES facility's profits obtained from the PCA-AP, AA, and MCS approaches, with their respective

TABLE IV: Comparison of PCA-AP, AA and MCS Approaches

Method		February			May			August			November		
		LB (\$)	UB (\$)	Burden	LB (\$)	UB (\$)	Burden	LB (\$)	UB (\$)	Burden	LB (\$)	UB (\$)	Burden
PCA-AP	95%	2,952	12,255	1min11s	8,428	13,567	6.37s	3,580	8,261	3min25s	4,631	8,306	33s
	97.5%	6,349	11,084	47s	8,089	13,077	7.47s	3,637	8,318	2min38s	4,075	8,673	29s
	99%	5,154	10,543	56s	8,133	12,991	26s	4,059	8,361	3min38s	4,020	8,429	41s
	100%	6,213	10,585	47s	8,086	12,924	41s	3,511	7,764	4min17s	4,400	8,367	46s
AA		1,971	12,673	12s	7,621	12,852	2.40s	2,505	8,885	48s	3,138	9,233	3.84s
MCS		5,560	8,240	1h27min	10,837	13,284	1h05min	5,978	8,105	1h20min	5,548	8,071	1h06min

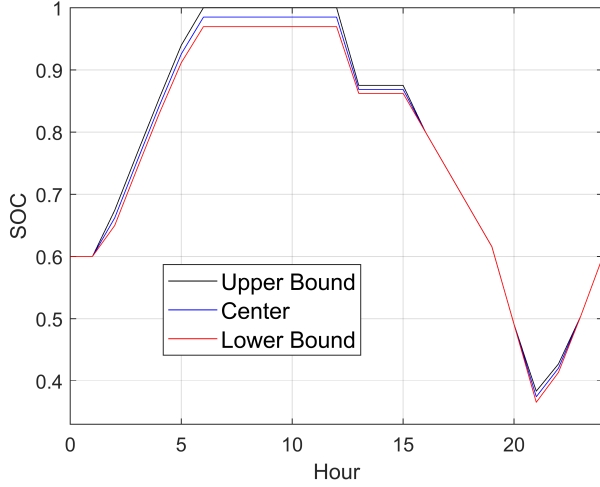


Fig. 6: SOC Profiles from AA Schedules for August.

computational burdens. Note that the AA approach presents the lowest computational burden, and its results envelops most of MCS and PCA-AP approaches with 99% and 100% information. However, it also results in the widest radius and pessimistic scenarios in all cases, which may yield significantly lower profits, due to the differences in the model with respect to the PCA-AP approach, and the better representation of the uncertainty intervals, as previously discussed. The PCA-AP approach with 95% and 97.5% information resulted in poor representations of the input data, in some cases, resulting in higher computational burden and profit intervals that differ significantly from the rest of the results. The PCA-AP results with 99% and 100% information envelop most of the MCS results.

Despite the PCA-AP model presents higher computational costs compared to the AA, in all cases, convergence was achieved in less than 5 minutes, rendering the proposed PCA-AP approach reasonable for day-ahead operations. The PCA-AP approach also presented less conservative results, i.e., tighter profit boundaries, yielding a more profitable margin of operation under pessimistic scenarios.

D. Concurrent Charging and Discharging

As previously stated in Section II, removing constraint (8) from the model enables the CAES facility to charge and discharge concurrently. Thus, to assess the facility's performance in this operational mode, the PCA-AP simulations were executed for the month of August.

TABLE V: Charging and/or discharging operation

PCA	August					
	Non-concurrent			Concurrent		
	Profit (\$)		Burden	Profit (\$)		Burden
95%	3,580	8,261	3min25s	-1,328	8,836	12min59s
97.5%	3,637	8,318	2min38s	1,868	8,510	16min56s
99%	4,059	8,361	3min38s	1,010	8,540	15min11s
100%	3,511	7,621	4min17s	665	8,045	22min22s

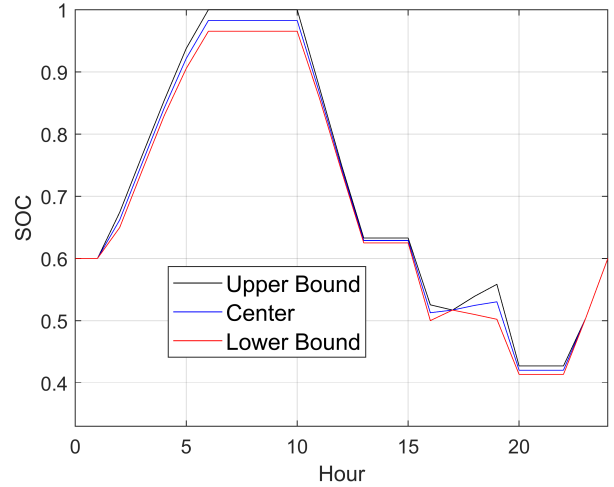


Fig. 7: Concurrent model SOC profile from PCA-AP 100% of information for August.

Table V presents a comparison with the previous model, where the facility operates in charging, discharging or idle modes, with the proposed concurrent charging/discharging model. The simultaneous charging and discharging allows more operational combinations and hence higher profits can be achieved. However, it also implies a greater computational cost and a wider profit interval.

Fig. 7 presents the optimum schedule for August for the PCA-AP with 100% of information for the concurrent model. Observe that the CAES facility is charging and discharging concurrently during hours 17-19, as reflected by the upper and lower bound moving in opposite directions in that period. In the non-concurrent model, there is a minimum charge/discharge due to the physical limitations of the CAES facility, while the concurrent model provides more flexibility since the total balance of air charged/discharged can be controlled. Thus, this model allows a smoother transition in the SOC profile.

V. CONCLUSIONS

This paper presented a novel PCA-AP self-scheduling model for a price-taker CAES facility operating with electricity price uncertainties. For an accurate representation, the thermodynamic characteristics of the CAES facility were considered in detail. Simulations were carried out considering different information levels, where a representation with lower information did not necessarily imply a lower computational burden, since poor representations of the input data impact the model performance.

The proposed PCA-AP approach was compared with an AA approach and validated with MCS. Despite the PCA-AP presenting a greater computational burden than the AA, it proved to be appropriate for day-ahead operations. In terms of profit, the PCA-AP approach yielded tighter intervals, thus resulting in more profitable margins of operation as compared to the AA approach. For validation purposes, both methods were benchmarked against the MCS approach, yielding lower computational costs while also enveloping the MCS results. This also demonstrated that, for the PCA-AP approach to be accurate, high PC information levels are required.

Finally, the possibility of the facility operating charging and discharging concurrently was tested and compared with the base model. Despite the greater computational cost, this approach achieved higher profits since the possible operational combinations of the facility increased. For future work, the authors intend to consider participation in the regulation market, thus expanding the profit opportunities for a CAES facility that can charge and discharge concurrently. Furthermore, the development of a price-maker model is also being considered.

APPENDIX A AA-BASED MODEL

This appendix presents the operational constraints of the AA-based model. Thus, the operational constraints (11)-(15) are represented in their affine form as follows, considering that continuous and not binary variables are represented in affine form:

$$P_{0,t}^C + \sum_{h=1}^3 P_{h,t}^C \leq \bar{P}^C x_t^C \quad \forall t \quad (76)$$

$$P_{0,t}^{SRC} + \sum_{h=1}^3 P_{h,t}^{SRC} \leq P_{0,t}^C - \sum_{h=1}^3 P_{h,t}^C - \underline{P}^C x_t^C \quad \forall t \quad (77)$$

$$P_{0,t}^{SRD} + \sum_{h=1}^3 P_{h,t}^{SRD} \leq \bar{P}^D x_t^D - P_{0,t}^D - \sum_{h=1}^3 P_{h,t}^D \quad \forall t \quad (78)$$

$$P_{0,t}^D - \sum_{h=1}^3 P_{h,t}^D \geq \underline{P}^D x_t^D \quad \forall t \quad (79)$$

$$0 \leq P_{0,t}^{ID} \pm \sum_{h=1}^3 P_{h,t}^{ID} \leq QSC x_t^{ID} \quad \forall t \quad (80)$$

The inter-temporal constraint (16) is presented as follows:

$$SOC_{0,t+1} = SOC_{0,t} + \frac{(air_{0,t}^C - air_{0,t}^D)3,600}{CA} \quad \forall t \quad (81)$$

$$\sum_{h=1}^p SOC_{h,t+1} = \sum_{h=1}^p SOC_{h,t} + \sum_{h=1}^p \left[\frac{(air_{h,t}^C - air_{h,t}^D)3,600}{CA} \right] \quad \forall t \quad (82)$$

$$SOC_{0,t+1} + \sum_{h=1}^p |SOC_{h,t+1}| \leq \overline{SOC} \quad \forall t \quad (83)$$

$$SOC_{0,t+1} - \sum_{h=1}^p |SOC_{h,t+1}| \geq \underline{SOC} \quad \forall t \quad (84)$$

$$SOC_{0,t} - \sum_{h=1}^p |SOC_{h,t}| \geq SOC^f \quad \forall t = T \quad (85)$$

The constraints for the mass of air charged (20)-(25) are presented in their affine forms as follows:

$$air_{0,t}^C = -0.3w_{0,t} + 1.95P_{0,t}^C \quad \forall t \quad (86)$$

$$\sum_{h=1}^3 air_{h,t}^C = \sum_{h=1}^3 [-0.3w_{h,t} + 1.95P_{h,t}^C] \quad \forall t \quad (87)$$

$$w_{0,t} - \sum_{h=1}^3 w_{h,t} \geq \underline{SOC} P_{0,t}^C + SOC_{0,t} \underline{P}^C - \underline{SOC} \underline{P}^C - M(1 - x_t^C) + \sum_{h=1}^3 [\underline{SOC} P_{h,t}^C + |SOC_{h,t}| \underline{P}^C] \quad \forall t \quad (88)$$

$$w_{0,t} - \sum_{h=1}^3 w_{h,t} \geq \overline{SOC} P_{0,t}^C + SOC_{0,t} \bar{P}^C - \overline{SOC} \bar{P}^C - M(1 - x_t^C) + \sum_{h=1}^3 [\overline{SOC} P_{h,t}^C + |SOC_{h,t}| \bar{P}^C] \quad \forall t \quad (89)$$

$$w_{0,t} + \sum_{h=1}^3 w_{h,t} \leq \underline{SOC} P_{0,t}^C + SOC_{0,t} \bar{P}^C - \underline{SOC} \bar{P}^C - M(1 - x_t^C) - \sum_{h=1}^3 [\underline{SOC} P_{h,t}^C + |SOC_{h,t}| \bar{P}^C] \quad \forall t \quad (90)$$

$$w_{0,t} + \sum_{h=1}^3 w_{h,t} \leq \overline{SOC} P_{0,t}^C + SOC_{0,t} \underline{P}^C - \overline{SOC} \underline{P}^C - M(1 - x_t^C) - \sum_{h=1}^3 [\overline{SOC} P_{h,t}^C + |SOC_{h,t}| \underline{P}^C] \quad \forall t \quad (91)$$

$$w_{0,t}, w_{h,t} \geq 0 \quad \forall t, h \quad (92)$$

The air discharged and cost of natural gas equations (28)-(32) are represented in their affine forms as follows:

$$P_{0,t}^D = \sum_{s=1}^2 (b_{t,s}^D + q_s u_{t,s}^D) \quad \forall t \quad (93)$$

$$P_{0,t}^D - \sum_{h=1}^3 P_{h,t}^D = \sum_{s'=1}^2 (b_{t,s'}^D + q_{s'} u_{t,s'}^D) \quad \forall t \quad (94)$$

$$\sum_{s=1}^2 u_{t,s}^D = x_t^D \quad \forall t \quad (95)$$

$$\sum_{s'=1}^2 u_{t,s'}^D = x_t^D \quad \forall t \quad (96)$$

$$0 \leq b_{t,s}^D \leq \bar{b}^D u_{t,s}^D \quad \forall t, s \quad (97)$$

$$0 \leq b_{t,s'}^D \leq \bar{b}^D u_{t,s'}^D \quad \forall t, s' \quad (98)$$

$$air_{0,t}^D = \sum_{s=1}^2 (ang_s^D b_{t,s}^D + \underline{D}_s u_{t,s}^D) \quad \forall t \quad (99)$$

$$\underline{air}_t^D = \sum_{s'=1}^2 (ang_{s'}^D b_{t,s'}^D + \underline{D}_{s'} u_{t,s'}^D) \quad \forall t \quad (100)$$

$$\sum_{h=1}^3 air_{h,t}^D = air_{0,t}^D - \underline{air}_t^D \quad \forall t \quad (101)$$

$$CO_{0,t}^{NG} = (53.57x_t^D + 4.215P_{0,t}^D)\pi^{NG} \quad \forall t \quad (102)$$

$$CO_{h,t}^{NG} = 4.215P_{h,t}^D\pi^{NG} \quad \forall t, h \quad (103)$$

APPENDIX B PCA-AP-BASED MODEL

This appendix presents the operational constraints of the PCA-AP-based model. Thus, for better representation, (64) can be rewritten as:

$$\hat{P}_t^X = P_{0,t}^X + \Delta P_t^X \quad \forall t, X \quad (104)$$

where:

$$\Delta P_t^X = \sum_Y \sum_{n=1}^{NPC} p_{n,t}^{X,Y} \Delta PC_n^Y \quad \forall t, X \quad (105)$$

The operational constraints (11)-(15) are reformulated as:

$$|\Delta P_t^X| = \sum_Y \sum_{n=1}^{NPC} |p_{n,t}^{X,Y}| |\overline{\Delta PC}_n^Y| \quad \forall t, X \quad (106)$$

$$P_{0,t}^C + |\Delta P_t^C| \leq \bar{P}^C x_t^C \quad \forall t \quad (107)$$

$$P_{0,t}^{SRC} + |\Delta P_t^{SRC}| \leq P_{0,t}^C - |\Delta P_t^C| - \underline{P}^C x_t^C \quad \forall t \quad (108)$$

$$P_{0,t}^{SRD} + |\Delta P_t^{SRD}| \leq \bar{P}^D x_t^D - P_{0,t}^D - |\Delta P_t^D| \quad \forall t \quad (109)$$

$$P_{0,t}^D - |\Delta P_t^D| \geq \underline{P}^D x_t^D \quad \forall t \quad (110)$$

$$0 \leq P_{0,t}^{ID} \pm |\Delta P_t^{ID}| \leq QSC x_t^{ID} \quad \forall t \quad (111)$$

The inter-temporal constraint (16) is presented as follows:

$$SOC_{0,t+1} = SOC_{0,t} + \frac{(air_{0,t}^C - air_{0,t}^D)3,600}{CA} \quad \forall t \quad (112)$$

$$\Delta SOC_{t+1} = \Delta SOC_t + \frac{(\Delta air_t^C - \Delta air_t^D)3,600}{CA} \quad \forall t \quad (113)$$

$$SOC_{0,t+1} + |\Delta SOC_t| \leq \overline{SOC} \quad \forall t \quad (114)$$

$$SOC_{0,t+1} - |\Delta SOC_t| \geq \underline{SOC} \quad \forall t \quad (115)$$

$$SOC_{0,t+1} - |\Delta SOC_t| \geq SOC^f \quad \forall t = T \quad (116)$$

The mass of air charged equations (20)-(25) are reformulated as:

$$air_{0,t}^C = -0.3w_{0,t} + 1.95P_{0,t}^C \quad \forall t \quad (117)$$

$$\Delta air_t^C = -\Delta w_t + 1.95 |\Delta P_t^C| \quad \forall t \quad (118)$$

$$w_{0,t} - \Delta w_t \geq \underline{SOC} P_{0,t}^C + SOC_{0,t} \underline{P}^C - \underline{SOC} \underline{P}^C - M(1 - x_t^C) + [\underline{SOC} |\Delta P_t^C| + |\Delta SOC_t| \underline{P}^C] \quad \forall t \quad (119)$$

$$w_{0,t} - \Delta w_t \geq \overline{SOC} P_{0,t}^C + SOC_{0,t} \bar{P}^C - \overline{SOC} \bar{P}^C - M(1 - x_t^C) + [\overline{SOC} |\Delta P_t^C| + |\Delta SOC_t| \bar{P}^C] \quad \forall t \quad (120)$$

$$w_{0,t} + \Delta w_t \leq \underline{SOC} P_{0,t}^C + SOC_{0,t} \underline{P}^C - \underline{SOC} \underline{P}^C - M(1 - x_t^C) - [\underline{SOC} |\Delta P_t^C| + |\Delta SOC_t| \underline{P}^C] \quad \forall t \quad (121)$$

$$w_{0,t} + \Delta w_t \leq \overline{SOC} P_{0,t}^C + SOC_{0,t} \bar{P}^C - \overline{SOC} \bar{P}^C - M(1 - x_t^C) - [\overline{SOC} |\Delta P_t^C| + |\Delta SOC_t| \bar{P}^C] \quad \forall t \quad (122)$$

$$w_{0,t}, \Delta w_t \geq 0 \quad \forall t, h \quad (123)$$

The air discharged and cost of natural gas equations (28)-(32) are reformulated as:

$$P_{0,t}^D = \sum_{s=1}^2 (b_{t,s}^D + \underline{q}_s u_{t,s}^D) \quad \forall t \quad (124)$$

$$P_{0,t}^D - |\Delta P_t^D| = \sum_{s'=1}^2 (b_{t,s'}^D + \underline{q}_{s'} u_{t,s'}^D) \quad \forall t \quad (125)$$

$$\sum_{s=1}^2 u_{t,s}^D = x_t^D \quad \forall t \quad (126)$$

$$\sum_{s'=1}^2 u_{t,s'}^D = x_t^D \quad \forall t \quad (127)$$

$$0 \leq b_{t,s}^D \leq \bar{b}^D u_{t,s}^D \quad \forall t, s \quad (128)$$

$$0 \leq b_{t,s'}^D \leq \bar{b}^D u_{t,s'}^D \quad \forall t, s' \quad (129)$$

$$air_{0,t}^D = \sum_{s=1}^2 (ang_s^D b_{t,s}^D + \underline{D}_s u_{t,s}^D) \quad \forall t \quad (130)$$

$$\underline{air}_t^D = \sum_{s'=1}^2 (ang_{s'}^D b_{t,s'}^D + \underline{D}_{s'} u_{t,s'}^D) \quad \forall t \quad (131)$$

$$\Delta air_t^D = air_{0,t}^D - \underline{air}_t^D \quad \forall t \quad (132)$$

$$CO_{0,t}^{NG} = (53.57x_t^D + 4.215P_{0,t}^D)\pi^{NG} \quad \forall t \quad (133)$$

$$\Delta CO_t^{NG} = 4.215\Delta P_t^D\pi^{NG} \quad \forall t, h \quad (134)$$

REFERENCES

- [1] G. T. Heydt, "Grand challenges in electric power engineering: extreme system reliability," *IEEE Power Engineering Society Summer Meeting*, vol. 3, pp. 1695–1697, July 2002.
- [2] B. Mohandes, M. S. E. Moursi, N. Hatziaargyriou, and S. E. Khatib, "A review of power system flexibility with high penetration of renewables," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3140–3155, 2019.
- [3] P. D. Lund, J. Lindgren, J. Mikkola, and J. Salpakari, "Review of energy system flexibility measures to enable high levels of variable renewable electricity," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 785–807, 2015.

- [4] J. Arteaga and H. Zareipour, "A price-maker/price-taker model for the operation of battery storage systems in electricity markets," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6912–6920, 2019.
- [5] H. Alharbi and K. Bhattacharya, "Participation of pumped hydro storage in energy and performance-based regulation markets," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4307–4323, 2020.
- [6] B. J. Davidson, I. Glendenning, R. D. Harman, A. B. Hart, B. J. Maddock, R. D. Moffitt, V. G. Newman, T. F. Smith, P. J. Worthington, and J. K. Wright, "Large-scale electrical energy storage," *IEE Proceedings A - Physical Science, Measurement and Instrumentation, Management and Education - Reviews*, vol. 127, no. 6, pp. 345–385, 1980.
- [7] T. V. Nguyen, "Integration of compressed air energy storage with wind turbine to provide energy source for combustion turbine generator," *IEEE PES Innovative Smart Grid Technologies, Europe*, pp. 1–5, 2014.
- [8] "Goderich A-CAES facility," Hydrostor, 2021. [Online]. Available: <https://www.hydrostor.ca/goderich-a-caes-facility/>
- [9] M. King, A. Jain, R. Bhakar, J. Mathur, and J. Wang, "Overview of current compressed air energy storage projects and analysis of the potential underground storage capacity in India and the UK," *Renewable and Sustainable Energy Reviews*, vol. 139, p. 110705, 2021.
- [10] I. Calero, C. A. Cañizares, and K. Bhattacharya, "Compressed air energy storage system modeling for power system studies," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3359–3371, 2019.
- [11] S. Shafiee, H. Zareipour, and A. M. Knight, "Considering thermodynamic characteristics of a CAES facility in self-scheduling in energy and reserve markets," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3476–3485, 2018.
- [12] M. F. Z. de Souza, C. A. Cañizares, and K. Bhattacharya, "Self-scheduling models of a CAES facility under uncertainties," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3607–3617, 2021.
- [13] J. D. Lara, D. E. Olivares, and C. A. Cañizares, "Robust energy management of isolated microgrids," *IEEE Systems Journal*, vol. 13, no. 1, pp. 680–691, March 2019.
- [14] M. Pirnia, C. A. Cañizares, K. Bhattacharya, and A. Vaccaro, "A novel affine arithmetic method to solve optimal power flow problems with uncertainties," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2775–2783, 2014.
- [15] D. Romero-Quete and C. A. Cañizares, "An affine arithmetic-based energy management system for isolated microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2989–2998, 2019.
- [16] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4742–4751, 2013.
- [17] A. Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32–51, 2016.
- [18] A. Vaccaro and C. A. Cañizares, "An affine arithmetic-based framework for uncertain power flow and optimal power flow studies," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 274–288, 2017.
- [19] A. Lorca and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1678–1688, 2017.
- [20] Y. Li, S. Miao, B. Yin, W. Yang, S. Zhang, X. Luo, and J. Wang, "A real-time dispatch model of CAES with considering the part-load characteristics and the power regulation uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 105, pp. 179 – 190, 2019.
- [21] J. Stolfi and L. H. D. Figueiredo, "Self-validated numerical methods and applications," *Trends in Applied Computational Mathematics*, vol. 4, no. 3, 2003.
- [22] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, 2016.
- [23] K. Molugaram and G. S. Rao, *Statistical Techniques for Transportation Engineering*. Butterworth-Heinemann, 2017.
- [24] W. L. Winston and M. A. Venkataramanan, *Operations research*. Thomson, 2003.
- [25] "Market prices," IESO. [Online]. Available: <http://reports.ieso.ca/public/PriceHOEPPredispOR/>
- [26] "Canada inflation," Statista. [Online]. Available: <https://www.statista.com/statistics/271247/inflation-rate-in-canada/>
- [27] R. Sakia, "The box-cox transformation technique: A review," *The Statistician*, vol. 41, 01 1992.
- [28] L. Kogan, "Small-sample inference and bootstrap," 2010. [Online]. Available: <https://ocw.mit.edu/courses/sloan-school-of-management/>

15-450-analytics-of-finance-fall-2010/lecture-notes/MIT15_450F10_lec09.pdf

- [29] H. S. Sandhu, L. Fang, and L. Guan, "Forecasting day-ahead price spikes for the Ontario electricity market," *Electric Power Systems Research*, vol. 141, pp. 450–459, 2016.



cused on compressed air energy storage participation in the electricity market, using robust optimization and/or affine arithmetic to model uncertainties.

Matheus Zambroni Received the diploma in electrical engineering from Universidade Federal de Itajuba (UNIFEI), Itajuba, Brazil, in 2015. Master in electrical engineering from UNIFEI in 2017. He obtained a PhD in Electrical Engineering at the University of Waterloo, ON, Canada, in 2022, and is currently a Research Scientist at NRCAN's CanmetENERGY in Varennes, QC, Canada. He had experience working with power systems stability during his undergrad, and power quality and signal processing during his master's. His PhD research focused



Engineer degree from the Escuela Politécnica Nacional (EPN) in Quito-Ecuador in 1984, where he held different academic and administrative positions between 1983 and 1993, and his MSc (1988) and PhD (1991) degrees in Electrical Engineering are from the University of Wisconsin-Madison. His research activities focus on the study of stability, control, optimization, modeling, simulation, and computational issues in bulk power systems, microgrids, and energy systems in the context of competitive energy markets and smart grids. In these areas, he has led or been an integral part of many grants and contracts from government agencies and private companies worth millions of dollars, and has collaborated with multiple industry and university researchers in Canada and abroad, supervising/co-supervising over 170 research fellows and graduate students. He has authored/co-authored multiple publications with high citation indices, including journal and conference papers, technical reports, book chapters, disclosures and patents, and has been invited to deliver keynote speeches, seminars, tutorials, and presentations at many institutions and conferences worldwide. He is the Editor-In-Chief of the IEEE Transactions on Smart Grid, the 2022-2023 IEEE Division VII Director of the IEEE and PES Boards, and a Fellow of the IEEE, a Fellow of the Royal Society of Canada, where he was the Director of the Applied Science and Engineering Division of the Academy of Science from 2017 to 2020, and a Fellow of the Canadian Academy of Engineering. He is also the recipient of the 2017 IEEE PES Outstanding Power Engineering Educator Award, the 2016 IEEE Canada Electric Power Medal, and of multiple IEEE PES Technical Council and Committee awards and recognitions, holding leadership positions in several IEEE-PES Committees, Working Groups, and Task Forces.

Claudio Cañizares (S'85–M'91–SM'00–F'07) is a University Professor and the Hydro One Endowed Chair at the Electrical and Computer Engineering (ECE) Department, and the Executive Director of the Waterloo Institute for Sustainable Energy (WISE) at the University of Waterloo, where he has held various academic and administrative positions since 1993 and has received multiple recognitions, especially the 2021-2022 Awards of Excellence in Graduate Supervision at both the University and Faculty of Engineering levels. He obtained the Electrical



Kankar Bhattacharya (M'95–SM'01–F'17) received the Ph.D. degree in electrical engineering from the Indian Institute of Technology Delhi, New Delhi, India, in 1993. He was with the Faculty of Indra Gandhi Institute of Development Research, Mumbai, India, from 1993 to 1998, and with the Department of Electric Power Engineering, Chalmers University of Technology, Gothenburg, Sweden, from 1998 to 2002. In 2003, he joined the Electrical and Computer Engineering (ECE) Department, University of Waterloo, Waterloo, ON,

Canada, where he is currently a Full Professor, University Chair, and the Chair of ECE. His current research interests include power system economics and operational aspects. He is a Registered Professional Engineer in the province of Ontario.



Álvaro Lorca (M'17) is an Assistant Professor in the Department of Electrical Engineering and the Department of Industrial and Systems Engineering at Pontificia Universidad Católica de Chile. He is also co-director of Vinken, a business unit at Dictuc dedicated to specialized consulting, advanced research, and technological developments for the energy industry. He received a Ph.D. in Operations Research from Georgia Institute of Technology, and an M.Sc. in Engineering and an Industrial Engineering degree with a diploma in Mathematical Engineering from

Pontificia Universidad Católica de Chile. His research and professional activities are centered in the application of mathematical methods and data science in energy planning, power system operations, renewable energy integration, smart grids, and resilience, among other topics. He has published in various specialized scientific journals, including Operations Research, European Journal of Operational Research, Production and Operations Management, IEEE Transactions on Power Systems, Energy, Energy Economics, Applied Energy, and Energy Policy, among others. He has worked with ISO New England, World Bank, Environmental Defense Fund, Coordinador Eléctrico Nacional, Comisión Nacional de Energía, Chilean Ministry of Energy, Transelec, and Fraunhofer Chile Research, among others. His research has been recognized with various awards, including the Best Paper in Energy Award from the INFORMS Section on Energy, Natural Resources, and the Environment, and the Best Paper Award from the IEEE Power and Energy Society.