IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE POWER & ENERGY SOCIETY SECTION

Received 12 September 2023, accepted 5 October 2023, date of publication 10 October 2023, date of current version 16 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3323594

RESEARCH ARTICLE

Machine Learning-Additional Decision Constraints for Improved MILP Day-Ahead Unit Commitment Method

MOHAMED IBRAHIM ABDELAZIZ SHEKEEW, (Graduate Student Member, IEEE), AND BALA VENKATESH, (Senior Member, IEEE)

Department of Electrical, Computer, and Biomedical Engineering, Toronto Metropolitan University, Toronto, ON M5B 2K3, Canada
Centre for Urban Energy, Toronto Metropolitan University, Toronto, ON M5B 2K3, Canada

Corresponding author: Bala Venkatesh (bala@torontomu.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant RGPIN-2019-05591; and in part by the Centre for Urban Energy, Toronto Metropolitan University, Toronto, Canada.

ABSTRACT This paper introduces a two-stage (offline and online) artificial neural network (ANN) driven constraint creator model to improve the computational quality of day-ahead unit commitment (DAUC) in power systems. The DAUC is crucial for planning 24-hour operations and complex bid clearing through mixed-integer linear programs (MILP). However, slow convergence is common due to system complexity. Machine learning (ML) based methods have been used to enhance MILP-DAUC. Nonetheless, they can lead to sub-optimality and infeasibility. To overcome these challenges, (1) this paper proposes in the offline stage the ANN-generators subset (AGS) that can predict part of the optimal MILP-DAUC decisions using an ANN model. Online, only ML-generated decisions of AGS are used to form the ANN-driven constraints to enhance the main MILP-DAUC, forming the proposed ANN-MILP-DAUC method. (2) A feasibility handling process is proposed to retain the infeasible ML states to be optimized by the main MILP-DAUC formulation. (3) The proposed model issues an artificial factor that provides the percentage of generators accurately predicted and used as an ML training performance metric. The ANN model was trained using optimal MILP-DAUC solutions. Test results on IEEE 14-bus and 118-bus systems reported solution time reductions of 61.43% and 70.1%, respectively. Larger Polish 2383-bus, 3012-bus, and Ontario systems reported time reductions in the range of 33% compared with the main MILP-DAUC method using MOSEK™, a commercial solver. No degradation in the optimal solution was observed for all test systems, and the proposed method provides a lower-objective solution for the same running time, leading to better solutions.

INDEX TERMS ANN-driven constraints generator, ANN-constrained unit commitment, mixed-integer linear programming.

NOMENCLATURE

A. ACRONYMS

DAUC	Day-ahead unit commitment.
ANN	Artificial neural network.
MILP-DAUC	Main MILP-DAUC formulation (1) -(20).
ANN-MILP-DAUC	ANN-Driven Constrained MILP-DAUC (1) – (20), (28) - (30).

B. INDICES AND SETS

n, NG	Index and set of generators, $n \in \{1, \ldots, NG\}$.
i, j, NB	Indices and set for bus number, $i, j \in \{1, \ldots, NB\}$.
m, NM	Index and set for segments, $m \in \{1, \ldots, NM\}$.
t, NT	Index and set for an hour, $t \in \{1 \ldots NT\}$.
k, NK	Index and set for branches, $k \in \{1, \ldots, NK\}$.
s, NS	Index and set of tested net load scenarios, $s \in \{1, \ldots, NS\}$.
Ω	ANN generators set, that provides error-free generator status.

The associate editor coordinating the review of this manuscript and approving it for publication was Ning Kang.

S_{ANN}^{ON}	Predictor of ANN-ON-driven constraints set (26).
S_{ANN}^{OFF}	Predictor of ANN-OFF-driven constraints set (27).

C. PARAMETERS

$\overline{PG_n}$, $\underline{PG_n}$	Max. /Min. output power of generator n.
$\overline{PG_{nm}}$ $\underline{PG_{nm}}$	Max. /Min. generating unit capacity of segment m.
$\overline{PL_K}$	The current carrying capacity of transmission line k.
bd_n, bs_n	Shutdown and startup costs, respectively.
br_n, ba_n, bb_{nm}	Reserve cost and fuel price coefficients.
x_k	The inductive impedance of branch k.
γ	Online hourly spinning reserve factor.
PD_{ti}	Net power demand at the i^{th} bus at hour t. Net power demand is defined as the difference between the actual load and the renewable generation.
$[B]$	Susceptance matrix of the power system.
$RG10_n$	10-minute spinning reserve capacity.
$RG60_n$	60-minute ramp-up/down limit.
MUT_n	Min uptime for generator n.
MDT_n	Min downtime for generator n.

D. VARIABLES

UD_{tn}, US_{tn}	Shutdown and startup binary variables.
UG_{tn}	Optimal UC schedule (Binary).
PG_{tnm}	Generator output power.
RG_{tn}	Spinning reserve of each unit.
SR_t	Hourly minimum system spinning reserve requirement.
PT_{ti}	Vector of hourly real power flow from the transmission lines to the connected bus i at time t.
δ_{ti}	Vector of bus voltage angles.

E. MATRICES

$\boldsymbol{PD}$	Hourly bus-wise net demand vector for systems scenarios NS; $\boldsymbol{PD} = \left[PD_{t,i}^{1} \cdots PD_{NTNB}^{NS}\right] \in \mathbb{R}^{(NB \cdot NT) \times NS}$.
$\boldsymbol{UG}$	Vector for the optimal generation schedules of all net load scenarios $\boldsymbol{PD}$, obtained using MILP-DAUC (1) – (20); $\boldsymbol{UG} = \left[UG_{t,n}^{1} \cdots UG_{NTNG}^{NS}\right] \in \{0, 1\}^{(NT \cdot NG) \times NS}$.
$\boldsymbol{Y}$	The ANN output vector corresponding to all hourly bus net load scenarios $\boldsymbol{PD}$ is given by matrix: $\boldsymbol{Y} = \left[Y_{t,n}^{1} \cdots Y_{NTNG}^{NS}\right] \in \{0, 1\}^{(NT \cdot NG) \times NS}$.
$\boldsymbol{E}$	The error matrix for the ANN output is defined as $\boldsymbol{E} \in \{0, 1\}^{(NT \cdot NG) \times NS}$.

I. INTRODUCTION

Day-ahead unit commitment (DAUC) is a crucial optimization problem in power system operations that involves technical and reliability constraints for successful operations. Practically, it is used for planning 24-hour operations and clearing complex bids. However, recent interest in detailed unit commitment constraints has allowed for mid-to-long-term planning, improving resource management strategies, maintenance scheduling, and hydrothermal coordination. Unit commitment also helps assess systems with operational flexibility challenges, such as variable renewables, emission caps, and less flexible generating units.

In 2022, Ontario's electricity demand increased by 2.8%, reaching 137.5 Terawatt hours (TWh), as a result of the province's economic recovery from the epidemic. In the next 20 years, a significant increase is expected in the net demand in Ontario, Canada, as shown in Fig. 1 [1].

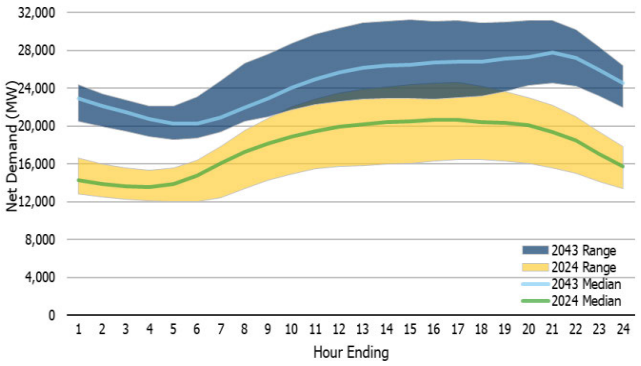

FIGURE 1. Forecasted increase in net demand of Ontario, Canada [1].

As electricity demand rises, power system operators need to ensure that there is sufficient generation capacity available to meet higher demand. This may require more generation units to be turned on and the selection process becomes more intricate. The increased number of units available for commitment and the potential need to bring additional units online quickly can increase the computational complexity of the optimization problem. Putting this DAUC in context, for Ontario's 38 GW system, which transacts energy worth $18 billion/year, a 1% improvement results in a $180 million benefit annually. This huge economic implication drives a constant desire to improve the DAUC process and to find efficiencies that benefit all stakeholders [2].

Computational Complexity -: The DAUC problem is formulated as a Mixed-Integer Linear Programming (MILP). Commercial solvers such as CPLEX and MOSEK were used to obtain the best-known operational costs [3]. However, MILP solvers exhibit slow convergence as binary variables increase with system complexity. Owing to its economic importance and computational complexity, the problem formulation and solution method of DAUC is an important and essential topic with serious practical implications and significant monetary impact for all power systems.

Typically, the system operator only has a short time to run the DAUC multiple times with different security scenarios every day (see Fig. 2 for the DAUC timeline of the

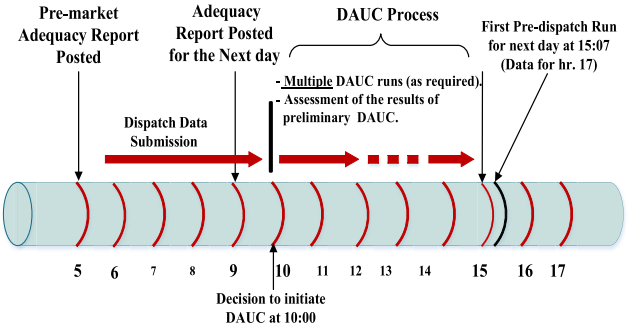

FIGURE 2. DAUC process timeline for IESO, Ontario, (Canada).

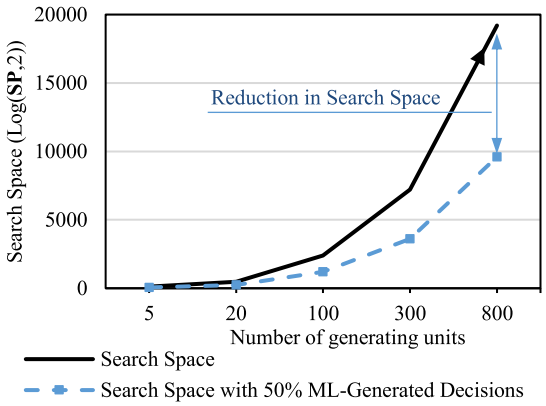

FIGURE 3. Search space reduction using 50% ML-Generated decisions.

Independent Electricity System Operator (IESO), Ontario, Canada) [2].

A. MILP SEARCH SPACE AND WORK OBJECTIVE

In the MILP, the search space refers to the set of all possible solutions that the optimization algorithm explores to find the optimal solution to the given problem. In the context of the DAUC process, the variables represent binary decisions such as whether a power generation unit is turned on (1) or off (0) during a certain time interval. In this case, the search space encompasses all possible combinations of on/off decisions for each unit over the given time horizon while ensuring that the total power generated meets the demand and respects the operational constraints of all units. Thus, the search space can be defined as $\boldsymbol{SP} = 2^{NG \cdot NT}$, where NG is the number of units and NT is the planning horizon [4], [5].

This study aims to introduce a set of Machine Learning (ML)-generated decision constraints for the main MILP-DAUC formulation to improve computational efficiency. For a well-trained ML model, these decisions are trusted to be a part of the optimal DAUC solution. Therefore, these ML constraints will significantly improve the computational quality, specifically the search space reduction. Figure 3 illustrates the effect of the additional constraints of 50% generation decisions using an ML model that equals $\boldsymbol{SP} = 2^{0.5 \cdot NG \cdot NT}$. It shows a significant reduction in the search space with an increase in the number of generating units in power systems.

Economic Loss-: Utilities need to determine the best achievable solution within a limited time so that dispatch instructions can be sent to generators. Using the latest MILP solvers, large-scale DAUC formulations can usually be solved with a higher relative Optimality Gap (OG) to meet the running time limits, where OG is the ratio of the difference between the best MILP solution and the continuous linear program (LP) solution divided by the continuous LP solution. However, it is impossible to achieve the best OG for large-scale systems within a limited time [6].

It is obvious that a further reduction in OG would yield a better solution and economic benefit for all stakeholders. Although MILP-based DAUC formulations provide a robust solution, little has been done to incorporate machine intelligence (MI) into these formulations to seek the least OG.

B. LITERATURE SURVEY

In the last few decades, various optimization techniques have been used to solve the DAUC challenge. These efforts have been directed toward improvements in conventional, intelligent, and hybrid optimization techniques. The literature on this topic is extensive and can be found in [6]. This study focused on the enhancement of the main MILP-DAUC formulation using ML models.

1) IMPROVEMENT IN MILP-DAUC FORMULATION

Recently, many efforts have been made to improve the performance of MILP-DAUC using tightening methods and constraint relaxation algorithms to obtain faster DAUC solutions.

Tightening methods: The DAUC formulation is approximated by an approach that considers a nonlinear objective function using perspective cuts. Slightly better solutions in less time were obtained, and the method was applied to systems that have up to 200 generating units, as in [7]. Tightened and compact DAUC formulations were presented for ramp rate constraints in [8], [9]. These formulations were tested using a Polish 2383-bus system. In [10], a formulation for the ramp rate constraint of gas turbines in a tight MILP UC formulation was presented. Another study presented the MILP formulation of stochastic DAUC with the chance constraint of ramping and operating reserves based on projected disjunctive reformulation [11]. In [12], transition variables were proposed to capture the intertemporal characteristics of the generating states between consecutive periods. A time-adaptive method was developed to reduce the computation time by aggregating the forecasted load periods and reducing the size of the DAUC problem [13].

Overall, the difficulties in obtaining the optimal solution for a large-scale DAUC formulation within the available time result in an increased OG, resulting in a suboptimal integer solution and economic loss. The review of MILP formulations for the DAUC challenge reveals that although these

methods provide improvements and are predominantly used in practical applications, a long computation time remains a challenge.

Constraint Relaxation Methods: These techniques involve temporarily relaxing or modifying specific constraints within the MILP formulation, thereby enabling quicker optimization. These relaxed constraints can be gradually tightened to obtain a feasible and nearly optimal solution. Notable research (e.g., [14]) has demonstrated substantial benefits, achieving significant reductions in solution times for large-scale power systems by selectively relaxing constraints. Similarly, another study [15] introduced a constraint relaxation technique tailored for renewable energy integration and demand response in power systems, reducing solution times while increasing flexibility to handle uncertainties that are particularly relevant to DAUC. Despite their potential to expedite solutions, constraint-relaxation methods have trade-offs. Relaxing constraints can yield initially infeasible or suboptimal solutions, necessitating additional computational effort to reestablish feasibility. Thus, selecting the right constraints for relaxation and determining the optimal relaxation levels remain complex tasks.

2) LITERATURE SURVEY – MACHINE LEARNING

Various recent studies have explored machine learning (ML) as a promising aid for solving power system operation and planning problems. In [16], a hybrid neural network with a simulated annealing approach was used to solve the DAUC optimization problem for a 10-unit system. The proposed approach uses an artificial neural network (ANN) to obtain a preschedule for the generation units based on the load profile as the input. Simulated annealing was then applied to solve DAUC. This method has been applied to a very simple system and does not consider transmission-network constraints.

A recent study used k-nearest neighbors (kNN) as a supervised ML method to train the total loads to predict unit commitment decisions [17]. Subsequently, they used a heuristic tuning method to select a particular percentage of the ML output to enhance the DAUC formulation. However, this method produces suboptimal or infeasible results because of ML prediction errors. In addition, the optimal tuning percentage cannot be guaranteed. In [18], the convolutional neural network (CNN) model was trained by wind power variations to predict all integer variables of the DAUC problem to convert the DAUC into the LP form. Despite the feasibility study that has been implemented, the DAUC solution cannot be trusted because the ML output may carry errors and be feasible simultaneously, leading to high-cost suboptimal solutions. In addition, this method assumes that loads are always constant. The stochastic UC challenge has been solved using ML's ability to predict the uncertainty of line disruptions caused by hurricanes [19]. Using historical data in [20] kNN was employed to eliminate unnecessary transmission constraints in the MILP-UC problem. Owing to prediction errors, some important transmission lines are removed from the formulation, which leads to either sub-optimality or infeasibility, as presented in the results for the 73-bus and 2000-bus systems. Eliminating redundant constraints (not only transmission constraints in UC) can be done easily in MILP solvers, such as in the presolve method of MOSEK presented in [21].

Table 1 shows the differences between the ML-based UC methods and the proposed method (PM).

TABLE 1. Comparison between ML-based UC methods.

	[16]	[17]	[18]	[22]	[PM]
DAUC objective	FC, SC	FC, SC SDC	FC, SC SDC	FC, SC, SDC, RC	FC, SC, SDC, RC
Network Constraints	X	√	X	√	√
Solution method	SA - ANN	MILP-kNN	MILP-CNN	MILP-ANN	MILP-ANN
System size	10 units	> 1800 bus	100 units	> 1800 bus	> 1800 bus
Model input	Total load	Total load	Wind curve	Net load	Hourly bus-wise net load
ML output Used for UC	Full	Partial	Full	Partial	Partial
Feasibility	X	X	X	√	√
Optimality	Sub-optimal	Sub-optimal	Sub-optimal	Near-Optimal	Optimal
ML variable reduction	√	√	√	√	X
ML-additional constraints	X	X	X	X	√
Presolve	X	X	X	√	√
Used ML errors	√	√	√	X	X

Note: Fuel Cost (FC), Startup Cost (SC), Shutdown Cost (SDC), Reserve Cost (RC).

In our recent study [22], ML was trained to shrink the DAUC formulation using a variable reduction method. In this study, ML prediction was used to create ANN-driven constraints, without shrinking the main DAUC formulation.

3) LITERATURE SURVEY – ML-BASED WARM START

The benefits of an ML-based warm start to solve the OPF problem were presented in [23], [24]. In [24], the ANN provided a prediction of natural gas pressures, which were used as the starting points for the Gas OPF problem. With the use of the warm start points, the results showed that the number of iterations and OG percentage were reduced. Importantly, for the large-scale DAUC problem, the ML-based warm-start solutions do not affect either the solution space or the solution method because modern MILP solvers already have advanced methods to find the best feasible starting point and condition the problem before starting, which renders these warm-start methods ineffective in comparison.

C. THE GAP IN ML-BASED DAUC METHODS

The literature presents recent efforts involving ML as a tool to enhance the MILP-DAUC. Various ML models have been explored, including hybrid neural networks

and simulated-annealing approaches [16]. However, these approaches have been applied to simplified systems and do not consider transmission-network constraints. Suboptimal results have also been reported when integrating 100% ML predictions into the DAUC formulation [17]. Challenges arise from ML prediction errors, feasibility concerns, and assumptions regarding load changes [18]. Furthermore, feasibility challenges persist in utilizing ML to eliminate transmission constraints in the MILP-UC problem [20].

D. CONTRIBUTIONS OF THIS PAPER

Unlike the ML-based UC literature, as presented in Table 1, the ML model is trained using the hourly net load at each bus to predict the UC decisions (not the total load, as in [16], [17]). Using only the total load change is not accurate because the effect of the hourly change in loads at each bus cannot be observed in the ML model. In addition, the proposed method provides additional ML constraints without shrinking the main MILP-DAUC formulation.

The main contributions of this work are summarized below:

1) Development of an ANN-driven constraint creator: An ANN-driven approach was created to accurately predict the trusted set of generators to be dispatched, yielding a set of ANN-driven constraints.
2) Enhanced Solution Efficiency: The incorporation of ANN-driven constraints into the MILP-DAUC method led to a remarkable reduction in solution times. This enhancement, achieved without any loss of optimality, resulted in solution times up to three times faster than the conventional MILP-DAUC formulation.
3) Feasibility Handling Process: To ensure the practical viability of our method, this process retains and refines the ML-generated decisions that may initially be unfeasible, allowing the main MILP-DAUC formulation to optimize them into the optimal solution.
4) Computational Quality: Tested cases, encompassing systems ranging from 14-bus to 3012-bus, consistently demonstrate that the proposed method converges to the same optimal solution as the MILP-DAUC method, but at a significantly accelerated pace. Furthermore, in larger systems, our approach yields lower-cost solutions within the same solution time, which ultimately results in substantial economic benefits.

The rest of this paper is organized as follows. Section II introduces model formulation and algorithms. Section III presents the methodology. Section IV presents system studies. Section V illustrates the proposed method's performance. The paper is concluded in Section VI.

II. ANN-MILP-DAUC FORMULATION

In this section, the ANN-MILP DAUC Model is described. First, the main MILP-DAUC formulation is presented. Thereafter, an ANN model is provided. This is followed by the development of predictor sets that created ANN-driven constraints.

A. MAIN DAUC MATHEMATICAL MODEL

The optimal solution of the DAUC provides the minimum total cost and is limited by the power system constraints, such as startup and downtimes of the generation units, ramp rates, capacity of transmission lines, and minimum hourly spinning reserves. The overall DAUC objective function includes the fuel costs of all types of generation, startup and shutdown costs, and reserve power costs. This objective is modeled as shown in (1).

$$\textit{Minimize}: \sum_{t=1}^{NT}\sum_{n=1}^{NG}\left[\begin{array}{l} bd_n \cdot UD_{tn} + bs_n \cdot US_{tn} + br_n \cdot RG_{tn} \\ +ba_n \cdot UG_{tn} + \sum_{m=1}^{NM} bb_{nm} \cdot PG_{tnm} \end{array}\right] \quad (1)$$

Subject to:

1) POWER FLOW CONSTRAINTS:

The transmission network power flow is examined in the DAUC using DC linear power flows, as shown in (2). In addition, the DC load flow approximates bus voltages equal to 1 p.u. with variable phase angles, as constrained in (3). Thus, the power flow for each line is defined in (4). Finally, the total injected power (i.e., generation power, demand, and power flow) is balanced at each bus for each period (one hour), as enforced by (5). PD_{ti} presents the net power demand at each bus at time t, and it is defined as the difference between the actual load (PD_{ti}^{act}) and renewable power generation (PG_{ti}^{Ren}) as shown in (5b).

$$[PT_{ti}] = \left[B'\right][\delta_{ti}]; \forall t, i \quad (2)$$

$$-\frac{\pi}{2} \le \delta_{ti} \le \frac{\pi}{2}; \forall t, i \quad (3)$$

$$\overline{PL_k} \le \left[\delta_{ti} - \delta_{tj}\right]/x_k \le \overline{PL_k}; \forall i, j \in k, t \quad (4)$$

$$\left[\sum_{n\in i}^{NG}\sum_{m}^{NM} PG_{tnm}\right] - PD_{ti} = PT_{ti}; \forall t, i \quad (5)$$

$$PD_{ti} = PD_{ti}^{act} - PG_{ti}^{Ren}; \forall t, i \quad (5b)$$

2) GENERATOR OUTPUT POWER LIMITS:

The generator output power limits are described by a set of constraints on the total power as a function of the unit status, segment-wise power output limits, and status of the units. The generation of each piecewise segment is limited to its maximum value, as shown in (6). Therefore, the generation power of each unit, which is the total power in each segment, is constrained by the thermal limit of the generating unit, as shown in (7).

$$0 \le PG_{tnm} \le \overline{PG}_{nm}; \forall t, n, m \quad (6)$$

$$UG_{tn} \cdot \underline{PG_n} \leq \left[PG_{tn} = \sum_{m}^{NM} PG_{tnm} \right] \leq UG_{tn} \cdot \overline{PG_n}; \forall t, n \tag{7}$$

3) SPINNING RESERVE CAPACITY:

The capacity of each generator to contribute to the spinning reserve is constrained as follows. The reserve power of each unit is limited by its reserve capacity in 10 min and the difference between its generation capacity and dispatch, as shown in (8). The online spinning reserve should be equal to the capacity of the largest source of the power system, as shown in (9). The online spinning reserve criterion is given by (10). When interconnected with other power systems, the parameter γ defines the required ratio, as is the case in Ontario, which lies within the Northeast Power Coordinating Council. It typically uses 0.25 for this value [25].

$$RG_{tn} \leq \left\{ \begin{matrix} UG_{tn} \cdot RG10_n, \\ \sum_{m}^{NM} \left(\overline{PG_{nm}} \cdot UG_{tn} - PG_{tnm} \right) \end{matrix} \right\}; \quad \forall t, n \tag{8}$$

$$SR_t \geq \left\{ UG_{tn} \cdot \left[\sum_{m}^{NM} \overline{PG_{nm}} \right] \right\}; \quad \forall t, n \tag{9}$$

$$\sum_{n}^{NG} RG_{tn} \geq \gamma . SR_t; \quad \forall t \tag{10}$$

4) GENERATORS RAMP RATES:

The committed generating units are constrained by their respective 60-minute ramp rate limits, as in (11).

$$-RG60_n \leq \sum_{m}^{NM} PG_{tnm} - \sum_{m}^{NM} PG_{t-1,nm} \leq RG60_n; \forall t, n \tag{11}$$

5) STARTUP AND SHUTDOWN CONSTRAINTS:

This set of constraints describes the unit startup, startup limits, and shutdown limits, as follows: The startup status of each unit is defined in (12), and the minimum startup limit for each unit is constrained in (13). The shutdown and minimum down limits for each turned-off unit are obtained in (14) and (15), respectively. The previous day commitment of each unit is considered in the DAUC formulation, as in (16) and (17).

$$US_{t+1,n} = max \left(UG_{t+1,n} - UG_{tn}, 0 \right); \forall t \in [0, \ldots, NT-1], n \tag{12}$$

$$US_{t+1,n} . MUT_n - \sum_{s=t+2}^{\min \left\{ \begin{matrix} NT, \\ t + MUT_n \end{matrix} \right\}} UG_{\text{sn}} \leq \max \left\{ \begin{matrix} 1, MUT_n \\ -NT + t \end{matrix} \right\}; \forall t \in [0, \ldots, NT-1], n \tag{13}$$

$$UD_{t+1,n} = max \left(UG_{tn} - UG_{t+1,n}, 0 \right); \forall t \in [0, \ldots, NT-1], n \tag{14}$$

$$\left(1 - UD_{t+1,n}\right) \cdot NT \geq \sum_{s=t+1}^{min \left\{ \begin{matrix} NT, \\ t + MDT_n \end{matrix} \right\}} UG_{sn}; \forall t \in [0, \ldots, NT-1]; n \tag{15}$$

$$If \left[IC_n > 0, MUT_n > IC_n \right], \sum_{t=1}^{MUT_n - IC_n} UG_{tn} \geq MUT_n - IC_n; \forall n \tag{16}$$

$$If \left[IC_n < 0, MUT_n > -IC_n \right], \sum_{t=1}^{MDT_n - IC_n} UG_{tn} \leq 0; \forall n \tag{17}$$

6) INTEGERS CONSTRAINTS:

Integer constraints for generator shutdown, startup, and status are imposed as shown in (18) – (20), respectively.

$$UD_{tn} \in \{0, 1\}; \forall t, n \tag{18}$$

$$US_{tn} \in \{0, 1\}; \forall t, n \tag{19}$$

$$UG_{tn} \in \{0, 1\}; \forall t, n \tag{20}$$

Equations (1) – (20) can be solved using a standard MILP solver. The optimal DAUC schedule is only UG_{tn} because the startup and shutdown vectors (US_{tn}, UD_{tn}) are mainly dependent on the unit status vector UG_{tn}.

B. DEVELOPMENT OF ANN-DRIVEN CONSTRAINTS CREATOR

The complex DAUC problem has a vast search space because it involves many intertemporal combinatorial constraints and a wide range of binary variables [5]. In this study, an ANN-driven constraint creator was devised to constrain subsets of generators to be turned on/off at certain hours. The MILP solver searches for a solution within the reduced search space. The ANN-driven constraint creator is described in the following subsections.

1) ANN-DRIVEN CONSTRAINTS CREATOR – ANN TRAINING DATA

In modern power systems, hourly bus net-load data and the corresponding DAUC schedules of generators are typically available from utilities in Ontario, Canada [26]. As detailed in Algorithm 1, an ANN model is trained using the hourly bus net load data as the input and the corresponding DAUC schedules of generators as the output. In the DAUC, for a period of NT hours, with NB buses, the number of input features for one load scenario would be equal $NB \times NT$ (PD_{ti}). This defines the number of nodes in the input layer. For the output layer, each unit status for each hour is represented, and hence, the number of output nodes equals $NG \times NT$ (UG_{tn}). Each day represented a scenario.

Using data from a few years, hundreds of scenarios NS are available for training, where the matrix of the hourly bus net

Algorithm 1 ANN-Training Model (Offline)

ANN-training model

1 If DAUC generator schedules are available: Use hourly bus net load scenarios ***PD*** and the corresponding generation schedule for all load scenarios ***UG***.
2 If DAUC generator schedules are not available: Solve the MILP-DAUC formulation from (1) – (20) for hourly bus net load scenarios ***PD*** and construct the corresponding generation schedule ***UG***.
3 Train a multi-layer feedforward backward propagation ANN model using ***PD*** as the input layer and ***UG*** as the output layer.
4 Use the resilient backpropagation (Rprop) training algorithm and mean squared error as a training performance metric.

load scenarios is given as $\boldsymbol{PD} = \left[PD^1_{t,i}, \cdots, PD^{NS}_{NTNB}\right] \in \mathbb{R}^{(NB\cdot NT)\times NS}$. The matrix of the generation schedule is $\boldsymbol{UG} = \left[UG^1_{t,n}, \cdots, UG^{NS}_{NTNG}\right] \in \{0,1\}^{(NT\cdot NG)\times NS}$. The ANN output corresponding to all hourly bus net load scenarios ***PD*** is given by matrix: $\boldsymbol{Y} = \left[Y^1_{t,n}, \cdots, Y^{NS}_{NTNG}\right] \in \{0,1\}^{(NT\cdot NG)\times NS}$. If the DAUC generator schedules are not available, MILP-DAUC methods (1) – (20) are solved for each net load scenario, and the generation schedule for each net load scenario is used as a target in the ANN training model.

2) ANN-DRIVEN CONSTRAINTS CREATOR – ANN STRUCTURE

Typically, an ANN model with three hidden layers was used to decrease the generalization error. Figure 4 shows the architecture of the ANN model. Moreover, the activation function in the hidden layers is a Hyperbolic Tangent Sigmoid (HTS) activation function, which is mathematically written as $F = \frac{e^x - e^{-x}}{e^{-x} + e^x}$, where F and x represent the neuron output and input, respectively, and the range of $F \in \{-1, .., 1\}$.

The relations between the input and output are described in (21) and (22), respectively:

$$H_i = F_i\left(PD^s_{tn} \cdot W^i + B^i\right) \tag{21}$$

$$Y^s_{tn} = F_h\left(W^h \cdot H_i + B^h\right) \tag{22}$$

where vector PD^s_{tn} is weighted by matrix W^i and adjusted by bias B^i and processed by the activation function $F(\cdot)$ in the input layer. The input layer yields vector H_i. Vector H is weighted by matrix W^h and adjusted by bias B^h and processed through $F(\cdot)$ in the hidden layer. The hidden layer yields an output Y^{ANN}_{tn}; this output is converted into binary, and when the output value is greater than zero, it is set to one, and zero otherwise.

a: ANN TRAINING ALGORITHM – OFFLINE

This work uses the Robust backpropagation algorithm (Rprop) as a training function to overcome system scalability (''squashing'' due to HTS transfer functions compressing an unlimited input range into a finite output range) and deliver high-quality training. The Rprop training algorithm is the fastest algorithm for pattern-recognition problems. Unlike the Levenberg–Marquardt algorithm, it requires a relatively small amount of memory to achieve the best training quality. Details of the Rprop algorithm are presented in [27].

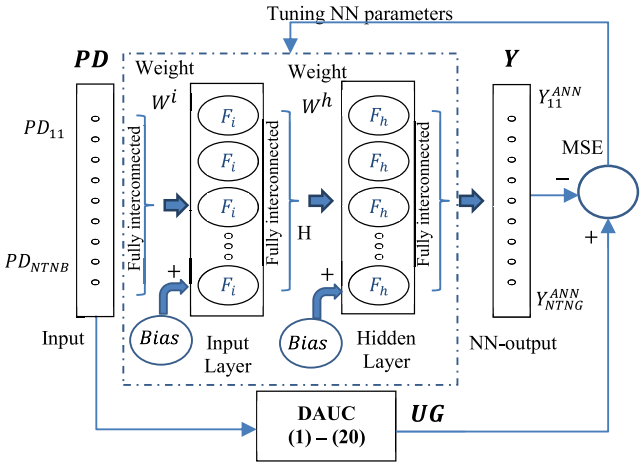

FIGURE 4. A three-layer neural network architecture.

b: ANN TUNING ALGORITHM

The ANN tuning parameters must be optimized to reduce the Mean Squared Error (MSE) between the estimated and actual system outputs. These parameters included the number of layers, amount of learning data, optimizer, batch size, number of neurons, activation function, and number of epochs. The number of layers is a highly sensitive factor in the precision of training. Overfitting occurred when too many layers were employed, whereas underfitting occurred when too few layers were used. These parameters are mostly empirical until they achieve their highest performance.

3) ANN-DRIVEN CONSTRAINTS CREATOR – OFFLINE ANN GENERATORS SET

Algorithm #2 illustrates the ANN output-selection method. Using Algorithm #1, the trained ANN model was used to predict the status of the generators ($UG^s_{t,n}$), a part of the MILP-DAUC solution. While the ANN training procedure minimizes errors, the nonlinear solution space of the DAUC precludes exact mapping by a trained ANN. Hence, although most generators are accurately predicted, a few are not. By examining the error plots of the ANN, the set of generators that can be accurately predicted can be ascertained. Accordingly, a method was presented to obtain an accurate part of the ANN output by examining the error matrix $\boldsymbol{E}$ of the ANN output.

The error matrix for the ANN output was defined as $\boldsymbol{E} \in \{0,1\}^{(NT\cdot NG)\times NS}$. The ANN Generators are the set of generators whose statuses are accurately predicted for load $\Omega \in \{0,1\}^{1\times NG}$. As the ANN model was constructed using

supervised learning, the error matrix $\boldsymbol{E}$ of the ANN output $\boldsymbol{Y}$ can be obtained as shown in (23).

$$\boldsymbol{E} = |\boldsymbol{UG} - \boldsymbol{Y}| \tag{23}$$

By analyzing $\boldsymbol{E}$, ANN generators are identified forming ANN-generators set Ω, as described in our previous work [22]. The ANN generators in this set were free of errors in all trained scenarios. In addition, the Artificial Factor ξ is defined as the percentage of ANN generators of the total generators that can be predicted accurately for all scenarios. This demonstrates the performance of the ANN. The ANN output for the generators of the ANN generator set Ω can be trusted to be accurate and can be used to create ANN-driven constraints without the possibility of infeasibility and degradation of the optimal solution from the regular MILP DAUC solution.

Algorithm 2 ANN Generator Set Ω (Offline)

1 Obtain the ANN output matrix $\boldsymbol{Y}$ for all load scenarios.

2 Use equation (23) to obtain the error matrix $\boldsymbol{E}$ for ANN output $\boldsymbol{Y}$.

3 **For** $n = 1 : NG$

3a Calculate the total error for each generator n in all trained scenarios as TE_n in (24).

$$TE_n = \sum_{s=1}^{NS} \sum_{t=1}^{NT} E\,(n) \tag{24}$$

3b **If** $TE_n = 0$, add generator n to the ANN Generators set $\boldsymbol{\Omega}$ that can be predicted with 100% accuracy for all scenarios. $\boldsymbol{\Omega_n} = 1$.
ELSE :$\Omega_{\boldsymbol{n}} = 0$.

4 Calculate the Artificial Factor ξ that presents the percentage of the total selected generators as follows:

$$\xi = \frac{1}{NG}\left[\sum_{n=1}^{NG} \Omega_n\right] \tag{25}$$

5 **If** $\xi \geq 0$
Export the ANN Generators Set: Ω.
Else
The training sets are not enough to generate robust ANN output.
Solution #1 – Increase training sets.
Solution #2 – Use a different training algorithm.

4) ANN-DRIVEN CONSTRAINTS CREATOR – PREDICTOR SETS CREATION - REAL-TIME

In real-time, the ANN model (from Algorithm #1) and ANN generator set Ω (from Algorithm #2) are combined to yield a predictor set of time and generator indices, where the generators are predicted to be ON or OFF accurately. Thereafter, ANN-driven constraints are developed using these sets and appended to the MILP-DAUC formulations (1) – (20).

Algorithm 3 MILP-DAUC Formulation With ANN-Driven Constraints Algorithm [Real-Time]

Stage 1: Construct ANN-driven Constraints Creator

1 Read input hourly bus net load data PD_{ti}.
2 Use ANN (trained in offline Algorithm #1) to determine the output of all generators.
3 Use ANN Generator Set Ω (completed in offline Algorithm #2) to construct predictor sets $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$ as in (26) and (27),
3a Check the feasibility for each generator $n \in \Omega$ using $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$ on security constraints (12) – (17).
3b If infeasible, then delete the decisions of n from predictor sets $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$.
4 Create ANN-driven constraints as in (28) – (30).

Stage 2: Enhance MILP-DAUC Formulation with ANN-Driven Constraints

1 Compose the MILP-DAUC formulation from (1) – (20) with ANN-driven constraints as in (28) – (30).
2 Solve the complete formulation with a standard MILP solver and issue commitment plus dispatch results.

Taking the ANN-generators set Ω and the output of the ANN model that considers the next day's net demand (PD_{ti}), two predictor sets, $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$ are created to define the status (ON/OFF) of each ANN generator, as shown in (26) and (27).

$$S^{ON}_{ANN} = \begin{cases} (t,n) \in S^{ON}_{ANN}; & if\ \{Y_{t,n} = 1, n \in \Omega\} \\ (t,n) \notin S^{ON}_{ANN}; & otherwise \end{cases} \tag{26}$$

$$S^{OFF}_{ANN} = \begin{cases} (t,n) \in S^{OFF}_{ANN} & if\ \{Y_{t,n} = 0, n \in \Omega\} \\ (t,n) \notin S^{OFF}_{ANN}; & otherwise \end{cases} \tag{27}$$

Accordingly, these two predictor sets, $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$ determine the ANN generators that are constrained (ON or OFF), forming ANN-driven constraints. Equations (28) – (30) model ANN-driven constraints.

$$UG_{tn} = 1; \forall\,(t,n) \in S^{ON}_{ANN} \tag{28}$$

$$UG_{tn} = 0; \forall\,(t,n) \in S^{OFF}_{ANN} \tag{29}$$

$$PG_{tnm} = 0; \forall m, (t,n) \in S^{OFF}_{ANN} \tag{30}$$

a: FEASIBILITY HANDLING PROCESS

A feasibility check must be applied to each predicted generator index of the ANN generator set. The predicted generator

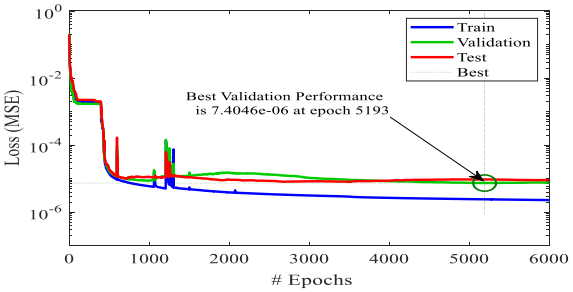

FIGURE 5. ANN training loss curve for IEEE 14-bus system.

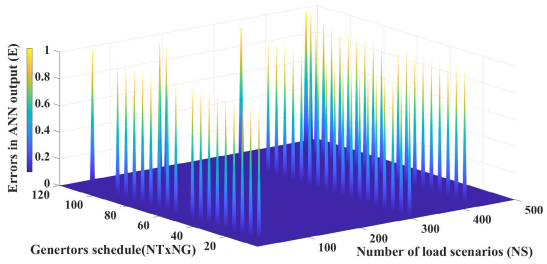

FIGURE 6. ANN-output errors E of all generators.

must be discarded from the predictor sets $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$, when it violates unit security constraints (12) – (17). Thus, the predictor sets $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$ are reduced because the infeasible decisions are converted to unknown states and optimized in the complete formulation, as in the following subsection.

b: COMPLETE ANN-MILP-DAUC ALGORITHM

The complete ANN-MILP-DAUC method is a combination of MILP-DAUC formulations (1) – (20) and ANN-driven constraints (26) – (30). These driven constraints are part of the optimal MILP-DAUC solution, and they have a direct effect on the search space reduction of the main formulation, as discussed in Subsection I-A. Equation (25) presents the percentage of ML-generated decisions. Thus, the search space reduction can be expressed as $\boldsymbol{SP} = 2^{\xi \cdot NG \cdot NT}$ which accelerates the MILP-DAUC solution speed.

Armed with the trained ANN model that drives constraint generation (26) – (30), the full set of steps for DAUC in real-time is shown in Algorithm 3.

C. ANN GENERATOR SET Ω ILLUSTRATIVE EXAMPLE

The IEEE 14-bus system has five generators. The ANN model was trained using 500 scenarios of 24-hour net load data and their corresponding 24-hour schedules of generators. The training loss curves are shown in Figure 5. As shown, the system was well-trained, with performance errors reaching 7.4E-06. Furthermore, the error matrix $\boldsymbol{E}$ for the ANN output of all load scenarios can be obtained using (23) and is plotted in Fig. 6. Upon examination of $\boldsymbol{E}$, it was discovered that the 2nd and 5th generators had errors. Thus, the ANN generator set Ω includes the 1st, 3rd, and 4th. Therefore: Ω = {1, 0, 1, 1, 0}. Furthermore, an Artificial Factor $\xi = 0.6$.

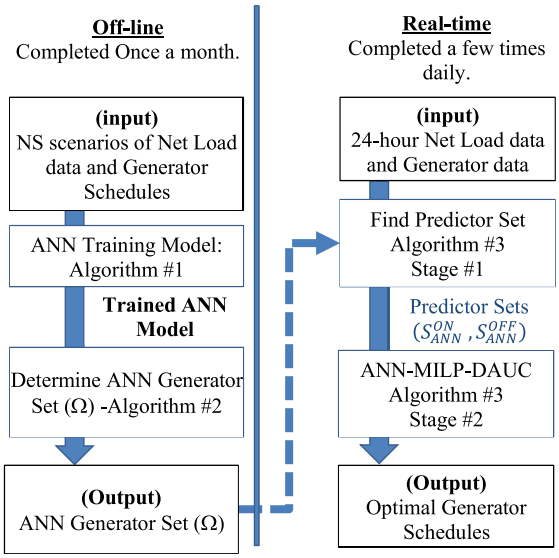

FIGURE 7. Complete ANN-MILP-DAUC method.

III. PROPOSED METHODOLOGY

The proposed method comprises offline and real-time components, as shown in Fig. 7. The offline part involves Algorithm 1 for ANN training and Algorithm 2 to determine the ANN generator set. The ANN model best maps the input 24-hour net load profile to the generator output. Conservatively, using a year's worth of hourly data should cover the entire training space and robustly enable ANN to drive constraint creation.

Furthermore, in many instances, a particular scenario may arise, which has not been seen before. In such cases, it is recommended to retrain the ANN model with updated data to ensure that the ANN remains current and evolves with the power system.

This ANN Generator Set Ω is fed into the real-time part of Algorithm #3. In Stage 1, the 24-hour day ahead demand is used with the trained ANN model to determine the generator outputs. This output with the ANN generator set Ω is used to determine the predictor sets $\left(S^{ON}_{ANN}, S^{OFF}_{ANN}\right)$ and create ANN-driven constraints. In Stage 2, the MILP-DAUC formulation is enhanced with ANN-driven constraints, forming the proposed ANN-MILP-DAUC method.

It is imperative to note that the ANN model only creates ANN-driven constraints to enhance the standard MILP-DAUC method. These ANN-driven constraints only reduce the search space and hence the solution time without causing infeasibility. The entire MILP formulation with these ANN-driven constraints remains robust and is uniquely solved by using MILP solvers.

IV. SYSTEMS STUDIES

The proposed ANN-MILP-DAUC method was implemented using MATLAB™ V. 9.7.012 (R2019B) and was solved using

TABLE 2. Benchmark systems data and ANN training parameters.

	IEEE 14-bus	IEEE 118-bus	Polish 2383-bus	Polish 3012-bus	Ontario Zonal System
Time horizon	24	24	24	24	24
Number of Generators	5	54	327	502	131
Size of the I/P layer $\boldsymbol{PD}$	336 × 500	2,832 × 148	57,192 × 198	72,288 × 200	360 × 208
Size of the O/P layer $\boldsymbol{UG}$	120 × 500	1,296 × 148	7,848 × 198	12,048 × 200	3,144 × 208
# Model Layers / # Neurons	3/60	3/80	3/80	3/80	3/80
# Learning Rate	0.01	0.01	0.01	0.01	0.01
Min. performance gradient	1e-60	1e-60	1e-60	1e-60	1e-60
# Epochs to train	6500	6500	9131	6500	6500
ANN Training Time (s) (Algorithm #1 & Algorithm #2)	300	1,200	4,500	6000	1000
ANN Training Accuracy (%)	99.9	99.8	97.5	97.5	98.47
ANN Testing Accuracy (%)	98.5	97.8	96.4	96.5	97.32
MAE	0.0011	0.0059	0.0117	0.0136	0.0075
RMSE	0.0332	0.0770	0.1081	0.1167	0.0864

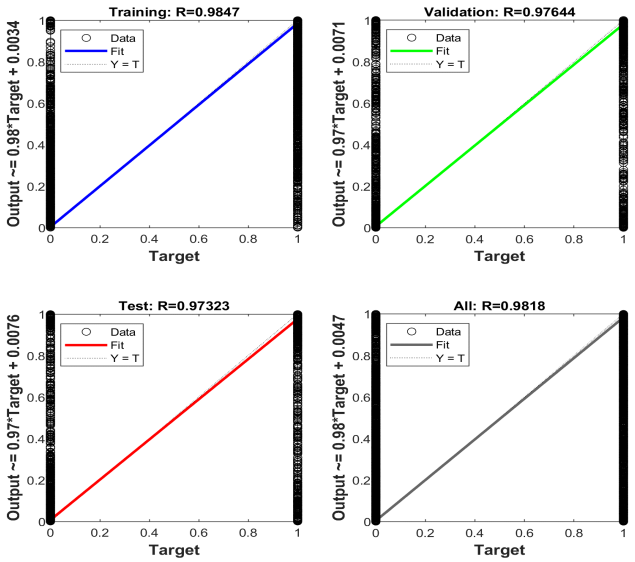

FIGURE 8. ANN-Training, testing, and validation accuracy for Ontario system.

TABLE 3. ANN-driven constraints generator results for the case studies.

	14-bus	118-bus	2383-bus	3012-bus	Ontario system
Artificial Generators set size Ω	3	29	160	247	54
Artificial Factor ξ	60%	54%	49%	49.2%	41%
# ANN ON constraints	72	216	504	1224	1272
# ANN OFF constraints	0	1,920	6,672	9408	48

the commercial solver MOSEK™ V 9.2.35 [28], both on a 3.60 GHz Intel™ Core processor i7 and 8 GB memory computer. Tests were executed for the proposed method on four benchmark systems: IEEE 14, 118-bus, Polish 2383-bus, and 3012-bus. In addition, the 10-zone Ontario, Canada network was tested. The results of all the completed tests are reported. The stopping criteria for the MILP solvers are commonly OG or maximum running time settings. In this section, for all systems, the stop time is set to 5000s, with varying solver OG settings relative to the scale of the system reported in Table 4. In addition, the MOSEK Presolve tool is used for all methods in this study to eliminate redundant constraints.

A. SYSTEMS DATA AND ANN TRAINING – ALGORITHM #1 [OFFLINE]

The data for the IEEE 14-bus system with five generators were obtained from [29]. The system has 24-hour time horizon loads at buses 2–6 and 9–14. The MILP-DAUC formulation for this system was characterized by 370 integer constraints, 520 continuous constraints, 120 integer variables, and 280 continuous variables. To train the ANN model, 500 day-ahead bus-wise net load scenarios $\boldsymbol{PD}$ were constructed as proposed in [22], and then the DAUC formulation in (1) – (20) was solved once for each of these load scenarios to obtain the corresponding generation commitment schedules $\boldsymbol{UG}$. These net load scenarios and their corresponding committed generating schedules were used as the input and target vectors, respectively, to train the ANN model. The IEEE 118-bus test system has 54 generators. The same training procedure as for the IEEE 14-bus system was used with 148 24-hour net load scenarios. Polish 2383-bus (327 units) and 3012-bus (502 units) data were collected from [29]. For each system, 198 scenarios and 200 scenario datasets with a 24-hour time horizon were created to train the ANN model. The training parameters and results of Algorithm 1 are listed in Table 2.

Furthermore, the proposed method was tested on a system in Ontario, Canada. The Ontario network is modeled as ten zones interconnected by 19 lines, including the interconnections with the neighboring networks (Quebec, Manitoba, Minnesota, Michigan, and New York), and 131 generation stations. A network diagram is presented in [30]. The system data were collected from the available online sources of IESO, the system operator of Ontario [31]. The MILP-DAUC formulation for this system was characterized by 12,600 integer constraints, 4,344 continuous constraints, 9,432 integer variables, and 6,672 continuous variables. The ANN model was trained using 208-scenario data sets using Algorithm #1.

Training Evaluation Metrics: The ANN training model was evaluated using two widely recognized error metrics: the Mean Average Error (MAE) and Root Mean Squared Error (RMSE). These metrics provide quantitative assessments of the accuracy and reliability of the training model, with lower values signifying a more effective and precise prediction methodology, ultimately reflecting a stronger alignment between the predicted and actual values. The following

TABLE 4. Average computational time (s) and average total cost ($) of random testing data for all case studies using the enhanced MILP-DAUC formulation with ANN-driven constraints (PM).

	IEEE 14-bus		IEEE 118-bus		Polish 2383-bus		Polish 3012-bus		Ontario Zonal system	
	PM	MILP-DAUC†	PM	MILP-DAUC†	PM	MILP-DAUC†	PM	MILP-DAUC†	PM	MILP-DAUC†
# of Testing Scenarios	500	500	148	148	198	198	200	200	208	208
# Integer constraints	576	504	5,904	5,208	35,256	31,416	126,910	120,982	13,896	12,600
# Continuous constraints	1,296	1,296	12,888	11,448	195,096	191,760	259,128	254,424	4,368	4,344
Integer Search Space of the DAUC	2^{48}	2^{120}	2^{600}	2^{1296}	2^{4008}	2^{7848}	2^{6120}	2^{12048}	2^{1848}	2^{3144}
Integer Search Space Reduction Factor	2^{72}	-	2^{696}	-	2^{3840}	-	2^{5928}	-	2^{1296}	-
Avg. CPU Time (s)	0.311	0.8063	26.86	89.82	37.96	56.2	129	192	7.349	10.966
Reduction in Time (%)	-61.43%	-	-70.1%	-	-32.5%	-	-32.8%	-	-33%	-
Average Total cost ($)	$1.254E+6	$1.254E+6	$1.199E+6	$1.199E+6	$2.716E+7	$2.717E+7	6.8447E+07	6.8451E+07	2.85E+07	2.85E+07
Average Change in Schedule Cost (%)	0.00%	-	0.00%	-	-0.022%	-	-0.007%	-	-6.5E-6%	-
Solver OG	1E-5	1E-5	1E-5	1E-5	0.006	0.006	0.006	0.006	1E-5	1E-5
Avg. converged OG	6E-6	6E-6	4E-6	5E-6	0.001455	0.001959	0.00348	0.00353	5E-7	1E-6

Note (†): MILP-DAUC solutions of the main formulation (1) – (20) of [17] and [18] have been averaged for the same tested data set.
PM: Proposed ANN-MILP-DAUC formulation.

equations describe these two error metrics and the training results are presented in Table 2. Furthermore, the ANN training accuracy for the Ontario system is graphed in Fig. 8.

$$MAE = \frac{1}{NS}\sum_{s=1}^{NS}\left|\left(\boldsymbol{UG}^{s} - \boldsymbol{Y}^{s}\right)\right| \tag{31}$$

$$RMSE = \sqrt{\frac{1}{NS}\sum_{s=1}^{NS}\left|\left(\boldsymbol{UG}^{s} - \boldsymbol{Y}^{s}\right)\right|^{2}} \tag{32}$$

B. ANN GENERATOR SET Ω – ALGORITHM #2 [OFFLINE]

After training the ANN model and achieving proper training performance, the ANN Generator Set Ω was created for each tested system using Algorithm #2. The ANN generators set Ω to provide zero errors in all ANN-trained scenarios. Table 3 lists the set sizes of each tested system. As shown, the ANN output selection method accurately predicted 60%, 54%, 49%, 49.2%, and 41% of the ANN generator output for IEEE 14-bus, IEEE 118-bus, Polish 2383-bus, Polish 3012-bus, and 10-zone Ontario systems, respectively.

C. TESTING THE ENHANCED MILP-DAUC FORMULATION WITH ANN-DRIVEN CONSTRAINTS – ALGORITHM #3

Using the ANN generator set Ω, Stage 1 of Algorithm 3 helps to create ANN-driven constraints. The details of these ANN-driven constraints are listed in Table 4 for all case studies.

The ANN-MILP-DAUC formulation enhanced with ANN-driven constraints, as presented in Stage 2 of Algorithm #3, was solved by MOSEK as a commercial MILP solver.

The effectiveness of the enhanced ANN-MILP-DAUC formulation with ANN-driven constraints was illustrated by comparing the results with the main MILP-DAUC formulation solutions (1) – (20) of [17] and [18]. The average computational quality and total cost for the tested datasets are listed in Table 4.

For the IEEE 14-bus system, the ANN model had an Artificial Factor ξ of 60%. This ANN generator set was used to generate 72 ANN-ON constraints (28), which were added to the ANN-MILP-DAUC formulation. The effect is shown by the total number of integer constraints, as listed in Table 4. As a result of these ANN-driven constraints, the enhanced formulation has a reduced search space, which leads to a 2.6 times faster solution (i.e., a 61.43% reduction in solution time), whereas the solution cost remains the same optimal value obtained using the main MILP-DAUC formulation without any change.

The test results of the proposed method on the IEEE 118-bus, Polish 2383-bus, and 3012-bus systems show significant benefits in terms of computation time without any degradation in the average optimal solution obtained by the main MILP-DAUC formulation and equal 70.1%, 32.5%, and 32.8% reductions in the computational time, respectively. In terms of the objective function, the proposed method showed improvement in Polish systems with average optimality gaps of 0.001455 and 0.00348, respectively, which resulted in an average cost reduction of −0.022% with a total benefit of $1.206 million in all tested data of the 2383-bus system, and −0.007% average cost reduction with total benefits $0.958 million for all data of the 3012-bus system.

It can be seen from the above results that the proposed ANN-MILP-DAUC method consistently achieved an equal or lower OG with time reductions for all cases. The total cost reduction is calculated using the following equation:

$$Totalbenefits\ (\$) = \frac{TC_{main} - TC_{PM}}{TC_{main}} \times NS \tag{33}$$

where TC_{main} is the average cost of the MILP-DAUC method and TC_{PM} is the average cost of the ANN-MILP-DAUC method for the number of tested load scenarios NS.

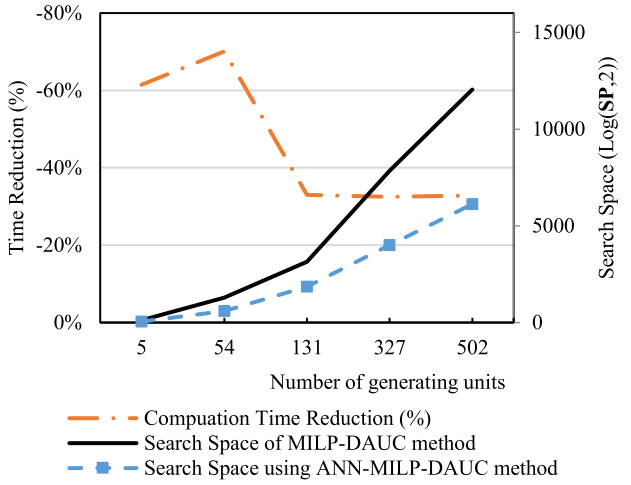

FIGURE 9. Effect of search space of ANN-MILP-DAUC model on solution time reduction.

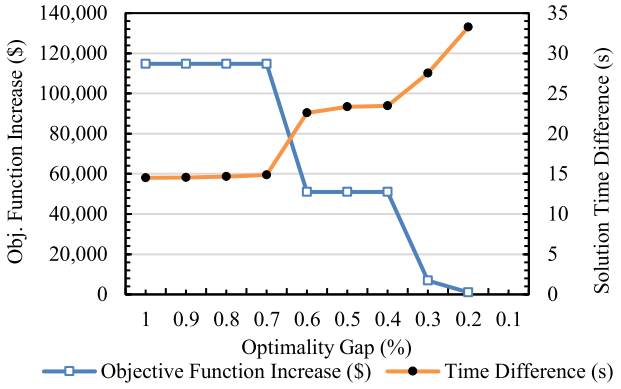

FIGURE 10. Effect of optimality gap for the Polish 2383-bus system on Objective Function and Solution Time.

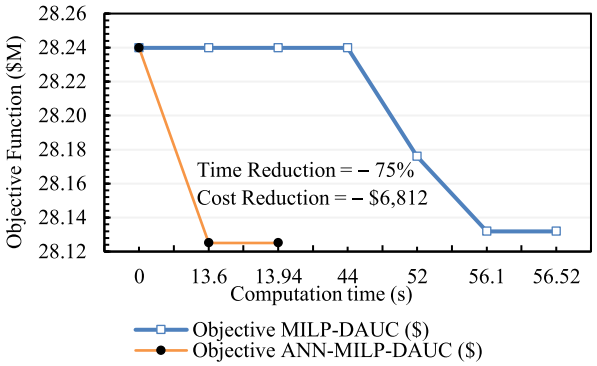

FIGURE 11. Solution evolution for Polish System using ANN-MILP-DAUC and MILP-DAUC methods (OG = 0.0024).

The proposed method narrows the solution search space. Table 4 shows the reduction in the solution space (the solution feasibility region). Objective function benefits are derived from the numerical efficiency of the MILP solver owing to the reduced size of the proposed ANN-MILP-DAUC method. Furthermore, these are derived from many tested samples, as mentioned in (33).

For the Ontario System, the results in Table 4 demonstrate the clear benefits of introducing ANN-driven constraints that enhance the MILP-DAUC formulation, leading to a reduced search space. It is 1.49 times faster (i.e., a 33% reduction in solution time), while the average solution cost is still robust, as the average cost obtained in the deterministic formulation is a marginal 6.5E-6% decrease. These results conclusively demonstrate that the proposed ANN-MILP-DAUC method is superior to the MILP-DAUC method, as it provides a significantly improved solution time for all tested systems, both small and large. Note that the objective function gains are zero for smaller systems and extremely minor for larger systems.

V. PERFORMANCE DISCUSSION

A. SEARCH SPACE AND COMPUTATION TIME

It is important to examine the effects of these new ANN-driven constraints in (26) – (30). They reduce the solution space by constraining the status variable UG, which models the status of the generators by limiting it to on or off. As discussed in Subsection I-A and Subsection II-B4, the search space reduction is expressed as $\boldsymbol{SP} = 2^{\xi \cdot NG \cdot NT}$, where ξ is the artificial factor.

For tested systems IEEE 14-bus, IEEE 118-bus, Ontario Polish 2383-bus, and 3012-bus systems, with an increase in artificial factors, the search space is reduced by a large order of 2^{72}, 2^{696}, 2^{1296}, 2^{3840}, and 2^{5928}, respectively. As presented in Table 4, these search space reductions enabled a significant reduction in solution time. Figure 9 shows the relationship between the search space and time reduction.

B. OPTIMALITY GAP, SOLUTION TIME, AND OPTIMAL SOLUTION

The optimality gap, optimal mixed-integer solution value, and solution time were interconnected. Two figures are presented to illustrate the benefits of the ANN-MILP-DAUC method over the MILP-DAUC method. Figure 10 shows the effect of the optimality gap for the Polish 2383-bus system. It may be reduced to provide a better optimal mixed-integer solution (a lower increase from the least-known solution) at a higher solution time. Hence, it is evident from Fig. 10 that more solution time is required when a better solution is sought. Furthermore, better solutions may not be reached owing to time restrictions.

Figure 11 shows the solution evolution of the ANN-MILP-DAUC and MILP-DAUC methods for the Polish 2383-bus system, considering an optimality gap of 0.0024 for a particular scenario.

It can be clearly seen that the ANN-MILP-DAUC method reaches the optimal solution faster than the conventional MILP-DAUC method owing to the reduced search space. The reduction in the solution time for this scenario was 75%, and

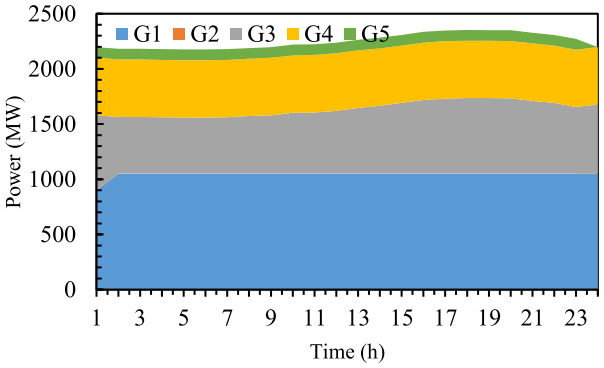

FIGURE 12. Identical DAUC dispatch of ANN-MILP-DAUC and MILP-DAUC methods for IEEE 14-bus.

TABLE 5. Performance of PM on UC states compared with the original formulation.

Systems		14-bus	118-bus	2383-bus	3012-bus	Ontario
Obj. Fn. (M$)	Main	1.0144	1.1707	28.1318	66.9226	30.0944
	PM	1.0144	1.1707	28.1250	66.8889	30.0944
Change (PM-Main) $		0	0	- 6,812	-33,700	0
# ON States	Main	95	683	1340	2555	1878
	PM	95	683	1080	2548	1878
# Different UC states (PM-Main)		0	0	- 260 (-19.4%)	-7 (-0.27%)	0
UC ON and OFF Similarities (%)		100	100	97	99	100

the reduction in the optimal solution was $6,812. Table 4 provides the average improvement over the 198 scenarios.

Combining the information in Fig. 10 and Fig. 11, it can be seen that the ANN-MILP-DAUC method provides the ability to seek a lower optimality gap, lower solution time, or a combination of both, as required. This allows for faster solutions and lower optimal costs, resulting in economic benefits.

C. PERFORMANCE OF ANN-MILP-DAUC ON UC SCHEDULES

It is important to examine the UC schedules generated from ANN-MILP-DAUC (PM) with the UC states of the original DAUC formulation (Main) (1) – (20). The UC results for a particular scenario within the dataset of the tested systems are listed in Table 5. The following conclusions can be drawn from the results:

1) For small-scale systems, that is, IEEE 14-bus, 118-bus, and Ontario zonal systems, the operating UC schedules (on states) are identical, with no discrepancies, and with the same operating cost. The UC dispatch results of the PM and MILP-DAUC are identical for the IEEE 14-bus system as shown in Table 5. Figure 12 shows the UC schedule for a particular day. As shown, generator # 2 is not economical and is always off, whereas unit #1 operates at full capacity throughout the day.

2) For large-scale systems, Polish 2383-bus and 3012-bus, the operating UC schedules are less than the schedules of the main formulation with full UC schedule (on and off) similarities of 97% and 99%, respectively, resulting in lower operating costs of $6,812 and $33,700 for both systems.

As discussed in Table 4, the search space of the PM is lower than that of the original UC method because of the added ANN constraints, which provide a faster solution and a better best-known MILP solution.

TABLE 6. Comparison with learning-based UC methods for the 118-bus system.

Learning Based- UC methods	Avg. Solver time %	Change in Obj. Fn	UC errors	ML-generated decision‡	Infeasibility percentage	Feasible instances
[16], [18]	-40%†	0.00%	3.40%	100%	65%	14
[17]	-58%†	0.00%	0.05%	70%	0.03%	39
[22]	-60%	0.00%	0.00%	50%	0.00%	40
Proposed Method	-70%	0.00%	0.00%	54%	0.00%	40

Note (†): Computational time for infeasible instances is taken as 2000 s. (‡) ML-generated decision refers to the percentage of ML decisions used to enhance the UC formulation to the total UC decisions of the main formulation.

D. PERFORMANCE WITH RECENT LEARNING-BASED METHODS

In this Subsection, the proposed method is compared with the most relevant learning-based UC methods in [16], [17], [18], and [22] to demonstrate its performance. In this comparison, the hourly bus-wise net demand was used as the input, and UC schedules were used as the output for the ANN model. Table 6 presents a comparison of 40 random datasets for the IEEE 118-bus. It was observed that [16] and [18] produced the lowest performance owing to the replacement of all integers of the UC formulation with the full ML-generated decisions that produced 3.40% prediction errors, resulting in a 65% infeasibility rate in the tested samples. Despite using a heuristic tuning method in [17] to use a certain portion of ML-generated decisions, the selected output contained 0.045% prediction errors, presenting an infeasibility rate of 0.02% in the dataset. Compared with our previous work [22], this model improves the solution speed for the IEEE 118 system by using a larger percentile of ML-generated decisions as driven constraints, instead of shrinking the main MILP-DAUC formulation.

For large-scale 3012 systems, the methods in [16] and [18] cannot solve the DAUC, as they provide a 100% infeasibility rate. For [17], it was infeasible in two out of 40 cases tested, whereas the proposed method provided feasible solutions and better computational quality for the IEEE 118-bus and Polish 3012-bus systems. Detailed results for the 3012-bus system are presented in Section IV.

It can be seen from the above results that the proposed method provides superior computation performance in terms of solution time and steadily provides a feasible and optimal solution, in contrast to [16], [17], [18] which are infeasible in some instances.

E. PRACTICAL CONSIDERATIONS

1) ASSESSMENT OF ANN OUTPUT ERROR

The reliability of the ANN output is rigorously validated using metrics like MAE and RMSE. These metrics quantify the accuracy of the ANN's predictions compared to actual MILP-DAUC decisions. Use of the proposed method should be sensitive to these parameters and strive to lower them. Additionally, the artificial factor can serve as an ML training performance metric. As discussed in Algorithm #2, if the artificial factor has a lower value, it indicates that the training datasets are insufficient for generating robust ANN output. To address this issue, consider two potential mitigation strategies: (1) increasing the size of the training sets and (2) exploring the use of a different training algorithm.

2) LOAD FORECAST UNCERTAINTY

To address load forecast uncertainty, the approach introduces robustness. Multiple sets of ANN-driven constraints are generated during the offline stage, each corresponding to different load scenarios. In the online stage, the set of constraints closest to the observed load is chosen, ensuring adaptability to varying load conditions. Hence, training on the varied data set and careful selection of Ω (ANN-generators set) shall ensure the adaptability of the method.

3) ANN TRAINING AND TUNING ALGORITHM

The training and tuning of the ANN model are pivotal for its effectiveness. In this approach, the Rprop algorithm is used to optimize the ANN's weights and biases. Rprop is particularly suited for our application due to its robustness in handling convergence and learning rate issues. The training dataset consists of historical data containing information on unit commitments, load, and other relevant variables. To prevent overfitting, techniques such as early stopping and regularization are judiciously applied. Furthermore, hyperparameter tuning is conducted, with grid search or Bayesian optimization being employed to optimize the ANN's architecture and parameters. Performance validation of the ANN is achieved through cross-validation, ensuring a high level of accuracy on the training data.

Effective tuning of ANN parameters is crucial for minimizing MSE. These parameters encompass the number of layers, epochs, training data, choice of optimizer, batch size, neuron configurations, and activation functions. The optimization of these metrics is primarily driven by empirical experimentation, fine-tuning them until optimal performance is attained, as detailed in Table 2.

VI. CONCLUSION

This study proposes a two-stage ANN-MILP-DAUC approach to enhance the MILP-DAUC computational challenges arising from power system scalability. To tackle these challenges effectively, this paper presents a multi-pronged solution: (1) Offline ANN-Guided Predictions: In offline stage, the concept of the ANN-generator subset (AGS) is introduced, achieved through an ANN model capable of predicting a portion of optimal MILP-DAUC decisions. This subset is a crucial component of the proposed approach. (2) Online Feasibility Handling: In the online phase, only the ML-generated decisions within the AGS are employed to construct ANN-driven constraints, enhancing the MILP-DAUC formulation and giving rise to the proposed ANN-MILP-DAUC method. (3) Artificial Factor and Performance Metric: This model introduces an artificial factor, serving as a metric that quantifies the percentage of generators accurately predicted, thus offering valuable insight into the performance of the ML model during training. The ANN model was trained using the optimal MILP-DAUC solutions, ensuring a strong foundation for decision-making.

The empirical validation of this approach on IEEE 14-bus and 118-bus systems demonstrated remarkable reductions in the solution times, reaching 61.43% and 70.1%, respectively. Moreover, larger systems, such as the Polish 2383-bus, 3012-bus, and real Ontario systems, report notable time reductions of 32.5%, 32.8%, and 33%, respectively, compared to the conventional MILP-DAUC method. Importantly, these improvements in computational efficiency do not compromise the optimality of the solutions, maintaining the optimal results within the same convergence tolerance. Furthermore, for larger systems, this method consistently delivers lower-objective solutions within equivalent solution times, thereby enhancing the overall computational quality.

REFERENCES

[1] Independent Electricity System Operator (IESO). (Dec. 2022). *Annual Planning Outlook Ontario's Electricity System Needs: 2024–2043*. Accessed: Sep. 1, 2023. [Online]. Available: https://www.ieso.ca/-/media/Files/IESO/Document-Library/planning-forecasts/apo/Dec2022/2022-Annual-Planning-Outlook.ashx

[2] IESO's Media Centre. *Electricity Data, Ontario, Canada*. Accessed: Sep. 1, 2023. [Online]. Available: https://www.ieso.ca/en/Corporate-IESO/Media/Overview

[3] B. Knueven, J. Ostrowski, and J.-P. Watson. *Mixed Integer Programming Formulations for the Unit Commitment Problem*. FERC. Accessed: Jun. 26, 2023. [Online]. Available: https://www.ferc.gov/sites/default/files/2020-09/W1-A-1-Knueven.pdf

[4] O. Guieu and J. W. Chinneck, "Analyzing infeasible mixed-integer and integer linear programs," *INFORMS J. Comput.*, vol. 11, no. 1, pp. 63–77, Feb. 1999, doi: 10.1287/ijoc.11.1.63.

[5] M. Rahmani, A. Kargarian, and G. Hug, "Search space reduction strategies for unit commitment problem," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Sep. 2016, pp. 1–5, doi: 10.1109/ISGT.2016.7781212.

[6] B. Saravanan, S. Das, S. Sikri, and D. P. Kothari, "A solution to the unit commitment problem—A review," *Frontiers Energy*, vol. 7, no. 2, pp. 223–236, Jun. 2013, doi: 10.1007/s11708-013-0240-3.

[7] A. Frangioni, C. Gentile, and F. Lacalandra, "Tighter approximated MILP formulations for unit commitment problems," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 105–113, Feb. 2009, doi: 10.1109/TPWRS.2008.2004744.

[8] G. Morales-España, J. M. Latorre, and A. Ramos, "Tight and compact MILP formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4897–4908, Nov. 2013, doi: 10.1109/TPWRS.2013.2251373.

[9] B. Yan, P. B. Luh, T. Zheng, D. A. Schiro, M. A. Bragin, F. Zhao, J. Zhao, and I. Lelic, "A systematic formulation tightening approach for unit commitment problems," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 782–794, Jan. 2020, doi: 10.1109/TPWRS.2019.2935003.

[10] K. Pan, Y. Guan, J.-P. Watson, and J. Wang, "Strengthened MILP formulation for certain gas turbine unit commitment problems," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1440–1448, Mar. 2016, doi: 10.1109/TPWRS.2015.2426139.
[11] W.-S. Tan and M. Shaaban, "A hybrid stochastic/deterministic unit commitment based on projected disjunctive MILP reformulation," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 5200–5201, Nov. 2016, doi: 10.1109/TPWRS.2016.2521326.
[12] S. Atakan, G. Lulli, and S. Sen, "A state transition MIP formulation for the unit commitment problem," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 736–748, Jan. 2018, doi: 10.1109/TPWRS.2017.2695964.
[13] S. Pineda, R. Fernández-Blanco, and J. M. Morales, "Time-adaptive unit commitment," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3869–3878, Sep. 2019, doi: 10.1109/TPWRS.2019.2903486.
[14] T. Ding, R. Bo, F. Li, Y. Gu, Q. Guo, and H. Sun, "Exact penalty function based constraint relaxation method for optimal power flow considering wind generation uncertainty," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1546–1547, May 2015, doi: 10.1109/TPWRS.2014.2341177.
[15] Z. Liang, Z. Dong, C. Li, C. Wu, and H. Chen, "A data-driven convex model for hybrid microgrid operation with bidirectional converters," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1313–1316, Mar. 2023, doi: 10.1109/TSG.2022.3193030.
[16] R. Nayak and J. D. Sharma, "A hybrid neural network and simulated annealing approach to the unit commitment problem," *Comput. Electr. Eng.*, vol. 26, no. 6, pp. 461–477, Aug. 2000, doi: 10.1016/S0045-7906(99)00037-3.
[17] Á. S. Xavier, F. Qiu, and S. Ahmed, "Learning to solve large-scale security-constrained unit commitment problems," *INFORMS J. Comput.*, vol. 33, no. 2, pp. 739–756, Oct. 2020, doi: 10.1287/ijoc.2020.0976.
[18] T. Wu, Y.-J. A. Zhang, and S. Wang, "Deep learning to optimize: Security-constrained unit commitment with uncertain wind power generation and BESSs," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 231–240, Jan. 2022, doi: 10.1109/TSTE.2021.3107848.
[19] F. Mohammadi, M. Sahraei-Ardakani, D. N. Trakas, and N. D. Hatziargyriou, "Machine learning assisted stochastic unit commitment during hurricanes with predictable line outages," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5131–5142, Nov. 2021, doi: 10.1109/TPWRS.2021.3069443.
[20] S. Pineda, J. M. Morales, and A. Jiménez-Cordero, "Data-driven screening of network constraints for unit commitment," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3695–3705, Sep. 2020, doi: 10.1109/TPWRS.2020.2980212.
[21] E. D. Andersen and K. D. Andersen, "Presolving in linear programming," *Math. Program.*, vol. 71, no. 2, pp. 221–245, Dec. 1995, doi: 10.1007/BF01586000.
[22] M. I. A. Shekeew and B. Venkatesh, "Learning-assisted variables reduction method for large-scale MILP unit commitment," *IEEE Open Access J. Power Energy*, vol. 10, pp. 245–258, 2023, doi: 10.1109/OAJPE.2023.3247989.
[23] R. Dobbe, O. Sondermeijer, D. Fridovich-Keil, D. Arnold, D. Callaway, and C. Tomlin, "Toward distributed energy services: Decentralizing optimal power flow with machine learning," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1296–1306, Mar. 2020, doi: 10.1109/TSG.2019.2935711.
[24] H. Liu, L. Yang, X. Shen, Q. Guo, H. Sun, and M. Shahidehpour, "A data-driven warm start approach for convex relaxation in optimal gas flow," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5948–5951, Nov. 2021, doi: 10.1109/TPWRS.2021.3107201.
[25] *NPCC Reserve Task Force on Coordination of Operations Revision Review Record, Northeast Power Coordinating Council Regional Reliability Reference Directory # 5 Reserve*. Accessed: Feb. 13, 2022. [Online]. Available: https://www.npcc.org/content/docs/public/program-areas/standards-and-criteria/regional-criteria/directories/directory-5-reserve-20200426.pdf
[26] IESO. *Historical Demand, Ontario, Canada*. Accessed: Sep. 10, 2023. [Online]. Available: https://www.ieso.ca/en/Power-Data/ Demand-Overview/Historical-Demand
[27] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," in *Proc. IEEE Int. Conf. Neural Netw.*, Mar. 1993, pp. 586–591, doi: 10.1109/ICNN.1993.298623.
[28] MOSEK-V9.2.35. (Jan. 2, 2021). *MOSEK ApS*. Accessed: Feb. 13, 2022. [Online]. Available: https://www.mosek.com/downloads/9.2.35/
[29] *Matpower Repository: IEEE 14-Bus, 118-Bus, and Polish 2383-Bus and 3012-Bus Systems Datasets*. Accessed: Feb. 18, 2023. [Online]. Available: https://matpower.org/docs/ref/matpower5.0/menu5.0.html
[30] J. Ma and B. Venkatesh, "Integrating net benefits test for demand response into optimal power flow formulation," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1362–1372, Mar. 2021, doi: 10.1109/TPWRS.2020.3020856.
[31] *IESO Public Reports: Zonal System Data*. IESO. Toronto, ON, Canada. Accessed: Sep. 11, 2023. [Online]. Available: http://reports.ieso.ca/public/

MOHAMED IBRAHIM ABDELAZIZ SHEKEEW (Graduate Student Member, IEEE) was born in Egypt. He received the B.Sc. and M.Sc. degrees (Hons.) in electrical engineering from Cairo University, Cairo, Egypt, in 2012 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Centre for Urban Energy, Toronto Metropolitan University, Toronto, ON, Canada. His research interests include power system optimization, energy storage, and renewable energy resources integration.

BALA VENKATESH (Senior Member, IEEE) received the Ph.D. degree from Anna University, India, in 2000. He is currently a Professor and the Academic Director of the Centre for Urban Energy, Toronto Metropolitan University, Toronto, Canada. His research interests include power system analysis and optimization. He is a registered Professional Engineer in the province of Ontario, Canada.