A Predictive Prescription Framework for Unit Commitment Using Boosting Ensemble Learning Algorithms

Ogun Yurdakul, *Student Member, IEEE,* Feng Qiu, *Senior Member, IEEE,* and Andreas Meyer

In this online companion, we present two appendices to the original manuscript. In Appendix A, we provide the detailed notation and the formulation of the CSUC problem employed in the publication. In Appendix B, we report the results on two additional case studies.

Appendix A
Notation and the Mathematical Formulation of the Conditional Stochastic Unit Commitment (CSUC) Problem

We begin by providing the nomenclature used in the publication.

Nomenclature

Sets,Indices

$\mathscr{H}, h$	set, index of hourly periods of a day
$\mathscr{N}, n$	set, index of nodes
$\mathscr{L}, \ell$	set, index of lines
$\mathscr{G}, g$	set, index of dispatchable generators (DGs)
$\mathscr{S}_g, s$	set, index of piecewise cost intervals for DG g

Variables

$u_g[h]$	$\in \{0,1\}$, commitment status of DG g in hour h
$v_g[h]$	$\in \{0,1\}$, startup status of DG g in hour h
$w_g[h]$	$\in \{0,1\}$, shutdown status of DG g in hour h
$p_g[h]$	power generated above minimum by DG g in hour h (MW)
$p_g^s[h]$	power from segment s for DG g in hour h (MW)
$p_c^n[h]$	curtailed load at node n in hour h (MW)

Parameters

Y/y	random net load/its observation
X/x	random auxiliary information/its observation
p_g°	initial power generated above minimum by DG g (MW)
u_g°	$\in \{0,1\}$, initial commitment status of DG g
$T_g^{\uparrow}(T_g^{\downarrow})$	minimum uptime (downtime) of DG g (h)
$T_g^{\uparrow,\circ}(T_g^{\downarrow,\circ})$	number of hours DG g has been online (offline) before the scheduling horizon (h)
$\bar{P}_g^s$	maximum power available from piecewise segment s for DG g (MW)
$\bar{P}_g(\underline{P}_g)$	maximum (minimum) power output of DG g (MW)
α_g^s	cost coefficient for piecewise segment s for DG g (\$/MWh)
α^c	penalty cost for load curtailment (\$/MWh)
α_g^v	startup cost of DG g (\$)
α_g^u	cost of DG g running and operating at $\underline{P}_g$ (\$/h)
$\Delta_g^{\uparrow}(\Delta_g^{\downarrow})$	ramp-up (ramp-down) rate limit of DG g (MW/h) q
$\Delta_g^{\uparrow,\circ}(\Delta_g^{\downarrow,\circ})$	startup (shutdown) rate limit of DG g (MW)
$\bar{f}_\ell(\underline{f}_\ell)$	maximum (minimum) real power flow allowed on line ℓ (MW)
Ψ_ℓ^n	injection shift factor of line ℓ with respect to node n

This work was supported in part by the U.S. Department of Energy Advanced Grid Modeling Program under Grant DE-OE0000875.

O. Yurdakul and F. Qiu are with the Energy Systems and Infrastructure Analysis Division, Argonne National Laboratory, Lemont, IL 60439 USA (e-mail: oyurdakul@anl.gov; fqiu@anl.gov).

A. Meyer is with the Department of Electrical Engineering and Computer Science, Technical University of Berlin, Berlin, Germany (andreas.meyer@tu-berlin.de).

The first-stage of the CSUC problem is expressed as

$$\underset{\substack{u_g[h], v_g[h],\\ w_g[h]}}{\text{minimize}} \quad \sum_{h\in\mathscr{H}}\sum_{g\in\mathscr{G}}\Big[\alpha_g^u u_g[h] + \alpha_g^v v_g[h]\Big] + \mathbb{E}\big[\mathcal{Q}(z,Y)|X=x\big], \tag{1}$$

subject to

$$u_g[h] - u_g[h-1] = v_g[h] - w_g[h] \qquad \forall h \in \mathscr{H}\setminus\{1\}, \tag{2}$$

$$u_g[h] - u_g^\circ = v_g[h] - w_g[h] \qquad \forall h \in \{1\}, \tag{3}$$

$$\sum_{h'=h-T_g^{\uparrow}+1}^{h} v_g[h'] \leq u_g[h] \qquad \forall h \in \{T_g^{\uparrow},\ldots,24\}, \tag{4}$$

$$\sum_{h'=h-T_g^{\downarrow}+1}^{h} w_g[h'] \leq 1 - u_g[h] \qquad \forall h \in \{T_g^{\downarrow},\ldots,24\}, \tag{5}$$

$$\sum_{h'=1}^{\min\{u_g^\circ\cdot(T_g^{\uparrow}-T_g^{\uparrow,\circ}),\,24\}} w_g[h'] = 0, \tag{6}$$

$$\sum_{h'=1}^{\min\{(1-u_g^\circ)\cdot(T_g^{\downarrow}-T_g^{\downarrow,\circ}),\,24\}} v_g[h'] = 0, \tag{7}$$

$$u_g[h], v_g[h], w_g[h] \in \{0,1\} \qquad \forall h \in \mathscr{H}, \tag{8}$$

where (2)–(8) hold for all DGs $g \in \mathscr{G}$. The objective (1) of the first stage is to minimize the commitment and startup costs plus the expected dispatch and load curtailment costs. We enforce by (2, 3) the logical constraints that relate the variables $u[h]$, $v[h]$, and $w[h]$, (4, 6) the minimum uptime, and (5, 7) the minimum downtime constraints. For a specific vector of first-stage decision variables z and an observation on net load $Y = \bar{y}$, the value function $\mathcal{Q}(z,\bar{y})$ is computed by solving the following second-stage problem:

$$\underset{\substack{p_g[h], p_g^s[h],\\ p_c^n[h]}}{\text{minimize}} \quad \sum_{h\in\mathscr{H}}\Bigg[\sum_{g\in\mathscr{G}}\sum_{s\in\mathscr{S}_g}\alpha_g^s p_g^s[h] + \sum_{n\in\mathscr{N}}\alpha_c p_c^n[h]\Bigg], \tag{9}$$

subject to

$$p_g[h] \le (\bar{P}_g - \underline{P}_g)u_g[h] \qquad \forall h \in \mathscr{H}, \tag{10}$$

$$p_g[h] \le p_g^\circ + \Delta_g^\uparrow u_g^\circ + (\Delta_g^{\uparrow,\circ} - \underline{P}_g)v_g[h] \qquad \forall h \in \{1\}, \tag{11}$$

$$p_g[h] \ge p_g^\circ - \Delta_g^\downarrow u_g^\circ - (\Delta_g^{\uparrow,\circ} - \Delta_g^\uparrow - \underline{P}_g)w_g[h] \quad \forall h \in \{1\}, \tag{12}$$

$$p_g[h] \le p_g[h-1] + \Delta_g^\uparrow u_g[h-1] + (\Delta_g^{\uparrow,\circ} - \underline{P}_g)v_g[h] \qquad \forall h \in \mathscr{H} \setminus \{1\}, \tag{13}$$

$$p_g[h] \ge p_g[h-1] - \Delta_g^\downarrow u_g[h-1] - (\Delta_g^{\uparrow,\circ} - \Delta_g^\uparrow - \underline{P}_g)w_g[h] \qquad \forall h \in \mathscr{H} \setminus \{1\}, \tag{14}$$

$$p_g[h] = \sum_{s \in \mathscr{S}_g} p_g^s[h] \qquad \forall h \in \mathscr{H}, \tag{15}$$

$$p_g^s[h] \le \bar{P}_g^s - \bar{P}_g^{s-1} \qquad \forall s \in \mathscr{S}_g, \forall h \in \mathscr{H}, \tag{16}$$

$$p_g[h] \in \mathbb{R}_+ \qquad \forall h \in \mathscr{H}, \tag{17}$$

$$p_g^s[h] \in \mathbb{R}_+ \qquad \forall s \in \mathscr{S}_g, \forall h \in \mathscr{H}, \tag{18}$$

$$p^n[h] = \sum_{g \in \mathscr{G}^n} p_g[h] + p_c^n[h] - \bar{y}_d^n[h] \quad \forall n \in \mathscr{N}, \forall h \in \mathscr{H}, \tag{19}$$

$$\sum_{n \in \mathscr{N}} p^n[h] = 0 \qquad \forall h \in \mathscr{H}, \tag{20}$$

$$\underline{f}_\ell \le \sum_{n \in \mathscr{N}} \Psi_\ell^n p^n[h] \le \bar{f}_\ell \qquad \forall \ell \in \mathscr{L}, \forall h \in \mathscr{H}, \tag{21}$$

$$p_c^n[h] \in \mathbb{R}_+ \qquad \forall n \in \mathscr{N}, \forall h \in \mathscr{H}, \tag{22}$$

where (10)–(22) hold for all DGs $g \in \mathscr{G}$. The objective (9) of the second-stage problem is to minimize the dispatch costs of DGs and the penalty cost incurred due to load curtailment. We enforce by (10)–(14) the generation and ramping limits based on the formulation laid out in [1]. The constraints on the power from each linear segment are stated in (15)–(16). We express the net real power injection at each node $n \in \mathscr{N}$ in (19) with the convention that $p^n[h] > 0$ if real power is injected into the system and state the system-wide power balance constraint in (20). We use the DC power flow model to state the transmission constraints and utilize injection shift factors (ISFs) for network representation [2]. In (21), we express the real power flow on each line ℓ in terms of nodal injections and ISFs and constrain it to be within its line flow limits. Finally, a nonnegativity constraint on $p_g[h]$, $p_g^s[h]$, and $p_c^n[h]$ is enforced (17), (18), and (22), respectively.

Appendix B
Additional Case Studies

We next report the results on two additional case studies conducted to further assess the performance of the proposed methodology.

A. Case Study I

In the first case study, we seek to investigate the performance of the proposed framework under different training set sizes D. To this end, we assess the out-of-sample performance of the prescriptions obtained under the following number of observations: $D \in \{3, 7, 14, 30, 61, 92\}$. The training set for each value of D is composed of the covariate and net load observations recorded on the first D days starting from April 1, 2018 (that is, for instance, the training set for $D = 3$ contains observations recorded from April 1 to April 3, 2018, and the training set for $D = 30$ contains the observations recorded in the entire month of April 2018). All experiments share the same test set constructed by using the 30 covariate and load observations recorded in April 2019.

For each investigated dataset size D, we train the utilized ML models (that is, AdaBoost, GBT, XGBoost, and RF) from scratch, and use each covariate observation in the test set to derive the weights assigned to historical observations and subsequently construct the prescription. Finally, we assess how the constructed prescription does out-of-sample given the actually materialized load observation. We repeat this procedure for all days in the test set and evaluate the total out-of-sample cost obtained under each ML model and training set size over all observations in the test set. In addition, for each prescription method and the training set size D, we compute the total time it takes to solve the problem and obtain the prescription.

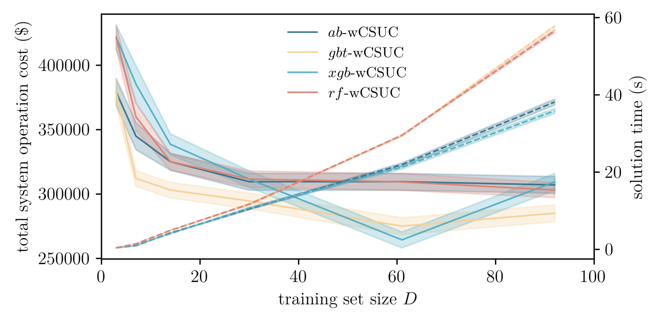

Fig. 1. Total system operation cost and the solution time of all proposed methods under different training set sizes D. The solid lines indicate the mean values (of both the total system operation cost and the solution time) over all 30 days in the test set. The shaded regions indicate one tenth of the standard deviation around the mean so as to avoid excessive overlap between the shaded regions.

We plot in Fig. 1 the average solution time and the total out-of-sample cost obtained under each method and training set size. The results in Fig. 1 show that the out-of-sample performance of the prescriptions obtained using the proposed method, by and large, improves (reduces) as the training set size increases. This notwithstanding, the out-of-sample performance exhibits diminishing returns, with D increasing beyond 30 making relatively insignificant contributions to the out-of-sample cost under most methods. In fact, under gbt-wCSUC and xgb-wCSUC, the out-of-sample cost rises as D increases from 61 to 92, occasioned by the inclusion of observations collected in June 2018 in the training set. We ascribe these results to the fact that the net load and covariate observations collected in June could have different underlying characteristics than those collected in April, which leads to a slight reduction in performance by hampering the quality of the prescriptions.

We further remark upon the tight coupling between the training set size and the solution time. Since we solve the extensive form of the CSUC problem, the solution time is directly influenced by the training set size. This brings to the fore the apparent trade-off between the solution time and the out-of-sample performance, which should especially be taken into account for $D \in \{30, 61, 92\}$, as increasing D beyond barely makes a dent in the out-of-sample performance, all the while significantly increasing the solution time.

References

[1] G. Morales-España, J. M. Latorre, and A. Ramos, “Tight and compact milp formulation for the thermal unit commitment problem,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4897–4908, 2013.

[2] K. E. Van Horn, A. D. Dominguez-Garcia, and P. W. Sauer, “Measurement-based real-time security-constrained economic dispatch,” *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3548–3560, 2015.