IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

IEEE POWER & ENERGY SOCIETY SECTION

Received 5 October 2023, accepted 21 October 2023, date of publication 25 October 2023, date of current version 1 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3327569

RESEARCH ARTICLE

Day-Ahead Scheduling of Wind-Hydro Balancing Group Operation to Maximize Expected Revenue Considering Wind Power Output Uncertainty

YU FUJIMOTO[1], (Member, IEEE), MAIKO INAGAKI[2], (Member, IEEE), AKIHISA KANEKO[2], (Member, IEEE), SHINICHIRO MINOTSU[3], AND YASUHIRO HAYASHI[2], (Member, IEEE)

[1]Advanced Collaborative Research Organization for Smart Society, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan
[2]Department of Electrical Engineering and Bioscience, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan
[3]Electric Power Development Company Ltd. (J-POWER), Chigasaki-shi, Kanagawa 253-0041, Japan

Corresponding author: Yu Fujimoto (yufujimoto@aoni.waseda.jp)

ABSTRACT To use energy generated in wind farms (WFs), which contain uncertainty in their output, as a primary power source in a power system, a sophisticated balancing operation scheme is required. This study focuses on a balancing group (BG) scheme combining WFs and a variable-speed pumped-storage hydro generator (PSHG), which has a large capacity to compensate for the WF output. The proposed BG operational scheduling approach aims to maximize the expected revenue obtained in the day-ahead power market under the uncertainty in WF output and market price by considering the operational constraints of the PSHG, i.e., the water storage capacity and frequency of operation for switching pump-up/-down. In such a BG scheme, it is crucial to consider time-varying and time-dependent uncertainties in WF output to manage PSHG capacity constraints, as well as to derive a reasonable plan in practical computation time. The proposed BG operational scheduling scheme derives a set of WF output scenarios that represents the heterogeneous and time-dependent characteristics of real-world WF output behavior from probability density distributions derived by a cutting-edge prediction approach and implements the expected revenue maximization problem with scenario-based chance constraints of water storage transition by introducing computationally effective iterative optimization algorithm based on surrogate functions. Simulation results suggest that the proposed scheme provides an effective BG operation by considering the uncertainty in the WF output.

INDEX TERMS Balancing group operation, wind power, pumped storage hydropower generation, probability density prediction, vine copula, expectation optimization, chance constraint, net-zero.

NOMENCLATURE LIST

SYMBOLS

α	Natural gradient descent vector used in Algorithm 1.
γ	$\in (0, 100)$. Percentile parameter for evaluation of PINAW and PICP.
ζ	$\in (0, 1)$. Threshold parameter for chance constraint.
η	Learning rate used in Algorithm 1.
$\boldsymbol{\theta}$	$= \{b_t, \boldsymbol{h}_t, g_t, l_t\}$. Decision variables of the BG scheduling problem.
κ	Operation efficiency of PSHG.
$\boldsymbol{\phi}$	Parameters of PDP for the target WF output.
ξ_t	$= \Pr(W_t < w_t)$. Cumulative lower-tail probability.
χ	Component model (regression tree) used in Algorithm 1.
ψ	Copula density function.
b_t	$\in \mathbb{R}$. Total BG output at timeslot t.

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

c_t Unit price of the day-ahead market at timeslot t.
$\underline{g}, \bar{g}$ The lower/upper limits of the PSHG generation during pump-down.
g_t $\in \mathbb{R}^+$. Generated output of PSHG under pumping-down operation at timeslot t.
$\boldsymbol{h}_t$ $= (h_t^g \in \{0, 1\}, h_t^l \in \{0, 1\})$. Pair of binary variables representing an operational status of PSHG at timeslot t; $h_t^g = 1$ and $h_t^l = 1$ indicate the PSHG is in pumping-down/-up operation, respectively.
i_t Unit price of imbalance charge at timeslot t.
$\underline{l}, \bar{l}$ The lower/upper limits of the PSHG load during pump-up.
$l(.)$ Log-likelihood function.
l_t $\in \mathbb{R}^+$. Load of PSHG under pumping-up operation at timeslot t.
r_t Revenue function for timeslot t.
w_t $\in [0, 1]$. Wind power generation [kWh/30min] of the target WF at timeslot t.
B Allowable operational state switching frequency per day.
$\bar{C}$ Storage capacity of the PSHG (i.e., maximum capacity of the upper reservoir).
$F(.)$ Objective function of the BG scheduling problem.
$\underline{F}(.), \bar{F}(.)$ Surrogate functions for lower/upper bounds of $F(.)$ used in Algorithm 2.
$I(.)$ The Fisher Information.
J The number of whole scenarios.
N^0 Initial partitioning parameter used in Algorithm 2.
N_t Partitioning parameter for timeslot t used in Algorithm 2.
$\bar{N}$ Partitioning limit parameter used in Algorithm 2.
S The number of scenarios adopted for PSHG operation scheduling.
$\mathcal{D}_t$ Historical dataset for learning model to predict the WF output of timeslot t.
$\mathcal{L}$ Index set of days included in the learning period.
$\mathcal{S}$ A set of scenarios for representing chance constraint of PSHG capacity.
$\mathcal{T}$ Half-hourly timeslots for WF output prediction target.
$\mathcal{T}_0$ $= \{1, \cdots, 48\}$. Half-hourly timeslots for the day-ahead market.
$\mathcal{T}_-$ Half-hourly timeslots for estimating initial water level.
$\mathcal{T}_+$ Half-hourly timeslots for optimization target.
$\cdot_t$ At the timeslot t (30 min time resolution).
$\hat{\cdot}$ Estimation/scheduling result.
$\cdot^*$ Realized value.
$[x]_+$ $= \max\{0, x\}$.

ABBREVIATIONS

BG Balancing group.
CWPSH Coordinated wind-pumped storage hydropower operation scheme proposed in [1].
EPDF Empirical probability density function.
EWF Ensemble weather forecast.
HWJB Hydro-wind joint bidding scheme proposed in [2].
JEPX Japan Electric Power eXchange.
NWP Numerical weather prediction.
PICP Prediction interval coverage probability.
PINAW Prediction interval normalized average width.
PSHG Pumped storage hydropower generation system.
PDP Probability density prediction.
SCADA Supervisory control and data acquisition system.
WF Wind farm.

I. INTRODUCTION

The massive demand for renewable energy sources towards realizing a net-zero society makes balancing electricity supply and demand difficult. Wind power suppliers are now focusing on enhancing the regular use of wind power output to maximize the potential of wind farms (WFs). In particular, power suppliers must bid their anticipated electricity supply values to a power market beforehand to provide electricity through the power grid; after the transaction is completed, a penalty cost is imposed when the actual supply deviates from the scheduled value. Therefore, to ensure the operational profitability of WFs, whose output is uncertain depending on wind conditions [3], operating WFs together with other dispatchable energy resources is desirable while utilizing output prediction results.

Currently, there are growing expectations for *balancing groups* (BGs) [4], [5], [6], in which such WFs and dispatchable energy resources are operated in a coordinated manner [7] as a group to achieve more systematic total generation planning and reduce penalty costs for market transactions. In the typical scheme, the BG operator submits the scheduled total BG power supply in advance, and during operation, the output of dispatchable energy resources is adjusted according to the WF output to reduce imbalance, i.e., the discrepancy between the scheduled and actual power generation; the revenue earned by the BG operator is determined based on the submitted output schedule, operational imbalance, electricity market prices, and prices for the imbalance charge. A key point to ensure the revenue in BG operation is considering the effect of uncertainty in WFs output during the scheduling phase. Many researchers have studied cooperative operation schemes of WFs with other dispatchable energy systems, such as battery storage systems [8], [9], [10], [11], [12], compressed air energy storage systems [13], [14], fuel cell systems [9], thermal power generators [15], [16], micro-turbines [9], and pumped-storage hydro generators (PSHGs) [2], [15], [17], [18], [19], [20], [21]. Among them, a framework that utilizes variable-speed

PSHGs [1] already installed in the current power system is an attractive option because it provides a large capacity to compensate for fluctuations in WFs output by adjusting the output under both pomp-up/-down operations without introducing additional energy resources.

A. LITERATURE REVIEW

Various research groups have studied frameworks to effectively realize BG operations combining PSHG and WF in market transactions. In particular, proper planning of PSHG operations in consideration of the uncertainty in wind power forecast results is vital to achieving sustainable operations; obviously, realizing WF output forecast with small errors has been addressed as one of the core technologies, but in addition to that, a scheme to derive a PSHG operation plan that minimizes the impact on BG operational results in case of forecast failure is essential.

For example, Parastegari et al. [19] have proposed a BG operation scheme based on stochastically generated scenarios, focusing on the universal empirical probability distribution of wind power output derived from historical data, to realize appropriate bidding for the market by a combination of WF and PSHG. Nazari et al. [21] have proposed using neural networks to predict wind power output deterministically for such cooperative operations; their framework introduced a single empirical distribution from the history of past forecast errors and used it in a heuristic optimization aimed at improving the profit, thus accounting for the effect of uncertainty in the forecast results. De La Nieta et al. [20] have derived a bidding strategy for the day-ahead market in a similar setting, with multiple scenarios of output for each timeslot that behave within the output range determined by the physical characteristics of the WF, and have proposed an operational problem for PSHG that maximizes profit by hedging risk in terms of conditional value at risk under the given scenario set. In the coordinated wind-pumped storage hydropower (CWPSH) operation framework proposed in [1], it is also assumed that a single probability density function represents uncertainty in the prediction error of the WF output. The usefulness of this profit maximization problem was discussed through a case study in which the uncertainty representation was applied to the actual history of the WF output. All of these frameworks assume *homogeneity* in wind power output uncertainty; in other words, these studies have not elaborately considered the statistical confidence level of the prediction results, which may vary with time and/or other factors.

On the other hand, for example, the study conducted by Varkani et al. [18] is one of the pioneering efforts to investigate a combined WF and PSHG operation scheme for participation in the day-ahead energy market; they have introduced a WF generation forecasting scheme based on neural networks similar to [21] but attempted to account for the effects of *heterogeneous* uncertainties which vary depending on the situation, by deriving the empirical probability distribution of the expected forecast error for each bidding timeslot separately. In a similarly pioneering effort, Garcia-Gonzalez et al. [17] have proposed to express the heterogeneity of uncertainty in the forecast results of WF output by leveraging multiple prediction results in BG operational planning for the day-ahead market. Cerejo et al. [2] have proposed the hydro-wind joint bidding (HWJB) framework using PSHG operational planning that takes into account the heterogeneity of WF output uncertainty by introducing an ad-hoc range of possible WF outputs under the assumption that it is proportional to the deterministic prediction results. He et al. [22] have made a similar argument in terms of the information gap decision theory [23]; they have considered dynamically varying ranges of possible WF outputs by introducing the similar ad-hoc assumption discussed in [2] to represent heterogeneity in the uncertainty. In recent years, real-world operators are gaining access to quasi-real-time information collected via supervisory control and data acquisition (SCADA) systems [24] and frequently updated wind forecast results based on numerical weather models [25]. In line with this, various cutting-edge frameworks have been proposed to realize interval prediction [26], [27], [28] and probability density prediction (PDP) [29], [30], [31], which flexibly represent the stochastic behavior of WF output that varies with local wind conditions for such BG operations [32].

Furthermore, the uncertainty in the forecast results of WF output is not only heterogeneous but also generally *time-dependent*; therefore, the forecast results for each timeslot of interest do not ideally behave independently [33]. Such time-dependent uncertainties in the prediction results make it even more difficult to appropriately estimate the energy capacity [34] needed to control energy storage (e.g., a reservoir in the PSHG) so that the BG output is consistent with the plan already submitted. Therefore, to achieve sustainable economic operation of the BG consisting of PSHG and WF based on bidding to the day-ahead market, it is necessary to advance BG operational management frameworks further to understand better the *heterogeneity* and *time-dependence* of wind power forecast results and to make the best use of them in operations. Table 1 summarizes how the heterogeneity and time-dependence of uncertainty in WF output have been assumed in the leading studies in the relevant field on the BG operation of WF and PSHG for the day-ahead market. This table suggests that while relevant studies have made a variety of assumptions about the statistical properties of WF output, there has been a particular lack of elaborate discussion of its temporal dependence.

B. AIMS AND CONTRIBUTIONS

This study proposes a schedule planning operation method for maximizing the expected profit from trading in the day-ahead market for BGs consisting of a WF and variable-speed PSHG. The proposed method utilizes the PDP [31], which provides rich probabilistic information, to predict the uncertainty in the WF output for the planning target period of

TABLE 1. Assumptions for uncertainty in WF output in relevant works.

Refs.	Assumptions for uncertainty in temporal behavior of WF output			
	Homogeneous	Heterogeneous	Independent	Dependent
Garcia-Gonzalez *et al.* (2008) [17]	-	✓	✓	-
Varkani *et al.* (2011) [18]	-	✓	✓	-
Parastegari *et al.* (2013) [19]	✓	-	✓	-
De la Nieta *et al.* (2016) [20]	✓	-	✓	-
Nazari *et al.* (2019) [21]	✓	-	✓	-
Cerejo *et al.* (2020) [2]	-	✓	✓	-
Li *et al.* (2021) [1]	✓	-	✓	-
Inagaki *et al.* (2021) [32]	-	✓	✓	-
He *et al.* (2022) [22]	-	✓	✓	-
Scheme proposed in this study.	-	✓	-	✓

the next day, and derives a schedule of the entire BG output plan and a schedule of the pump-down/-up operation of the PSHG. A feature of the proposed method is that it explicitly models the time-dependence of the wind power output distribution over the planning target period using vine copula [35], based on the historical sequences of the wind power output and those of past PDP results; the PDP approach, which consists of the gradient boosting concept [36] that has been discussed in the field of machine learning, allows for a precise representation of the heterogeneity of WF output uncertainty, and modeling with vine copula allows its time dependency to be incorporated appropriately into the BG operational management methodology. The proposed framework provides plausible scenarios for the transition of water level in the reservoir determined by the accumulation of PSHG control results over the planning period, thereby introducing additional appropriate chance constraints on the PSHG water storage for the optimal planning problem. The chance-constrained expectation optimization problem formulated in this study can be solved by iterative discrete bounding approximations [37] to efficiently evaluate the tight upper/lower bounds of the objective function in a computationally efficient manner; by applying these ideas, the proposed method enables the derivation of the BG operation plan within several hours for decision-making from the derivation of the WF output forecast results to the submission of the BG plan to the day-ahead market. The effectiveness of the proposed method was evaluated in terms of realized revenues through numerical simulations using actual output data measured at a WF site located in eastern Japan, market prices, and imbalance charges measured over a year. The contributions of this study are summarized as follows.

- We formulate a day-ahead wind-hydro BG scheduling problem for maximizing expected revenue, particularly considering uncertainty in wind power output.
- We suggest using a cutting-edge PDP scheme to account for the impact of heterogeneous and time-varying uncertainty in WF output in the expected revenue maximization problem.
- We propose a scenario-based PSHG capacity chance-constraints scheme by modeling statistical time dependence to address the impact of uncertainty in WF output behavior.
- We propose an efficient iterative optimization approach using surrogate functions representing the upper and lower bounds of the objective function which is difficult to optimize directly.

Numerical results were evaluated considering the effect of system coordination in the BG setting and the impacts of considering heterogeneity and time dependence in the WF output distribution.

C. PAPER ORGANIZATION

The remainder of this manuscript is organized as follows. Section II introduces the framework of BG operation, which consists of a WF and PSHG. In Section III, we formulate the BG operation scheduling problem regarding the expectation optimization problem and propose an approach to solve it efficiently. Section IV describes the simulation experiments conducted to evaluate the usefulness of the proposed method. Finally, Section V provides a concluding remark.

II. BALANCING GROUP OPERATION

We focus on the day-ahead market scheme launched in the Japan Electric Power eXchange (JEPX)[1] in 2019, which requires plans to be submitted by 10:00 on the previous day for 48 half-hourly timeslots for the next day. We will denote by $\mathcal{T}_0 = \{1, \ldots, 48\}$ the 30 min timeslots of a daily operation. Let w_t be the wind power generation of the target WF at timeslot $t \in \mathcal{T}_0$ and $g_t \in \mathbb{R}^+$ and $l_t \in \mathbb{R}^+$ be the outputs of the PSHG under the pump-down/-up operation, respectively. In addition, let $\boldsymbol{h}_t = (h_t^g \in \{0, 1\}, h_t^l \in \{0, 1\})$ be a pair of binary variables that represent the operational status (pumped-down/-up) of PSHG at the corresponding timing. The output of the WF, w_t, can behave unexpectedly depending on the weather, ensuring that the power is generated as scheduled difficult. Therefore, in this study, we assume that the BG operation is subject to the direct scheduling of the total system output,

$$b_t = w_t + h_t^g g_t - h_t^l l_t, \tag{1}$$

to compensate for the uncertainty by leveraging the controllability of the PSHG. Under the given BG scheduling results, $(\hat{b}_t, \hat{\boldsymbol{h}}_t)$, the revenue obtained for this timeslot via the

[1] "Japan Electric Power eXchange," [Online] Available: https://www.jepex.jp/en/, accessed: 2023-08-02.

day-ahead market can be written as:

$$r_t(w_t) = c_t\hat{b}_t - i_t\left[\hat{b}_t - (w_t + \hat{h}_t^g g_t^* - \hat{h}_t^l l_t^*)\right]_+, \quad (2)$$

where c_t and i_t are the unit values of the day-ahead market price and imbalance charge, respectively, and $[x]_+ = \max\{0, x\}$. The first term in Eq. (2) indicates that bidding on the BG operation plan, $\hat{b}_t$, in the day-ahead market for timeslot t in the scheduling phase results in a fundamental revenue based on the unit price. Meanwhile, the second term indicates the imbalance penalty based on the unit cost determined from the market balance at the timing of the actual transaction; the deviation between the scheduled BG operation $\hat{b}_t$ and BG output realized by the actual output, g_t^* and l_t^* of the PSHG, which operates to compensate for deviations from the assumed wind power output under system operational constraints, is settled according to this form. Note that the second term implies that when the actual BG output is likely exceeded the planned value $\hat{b}_t$ in timeslot t, the imbalance can be avoided via wind power curtailment.

The output, g_t^* and l_t^*, of the PSHG that defines the final BG output at timeslot t in Eq. (2) is assumed to be given by the following control result under the actual WF output w_t of the target day:

$$(g_t^*, l_t^*) = \underset{(g_t, l_t)}{\operatorname{argmin}} \left\{\hat{b}_t - (w_t + h_t^{g*} g_t - h_t^{l*} l_t)\right\} \quad (t \in \mathcal{T}_0), \quad (3)$$

subject to

$$\underline{g} \le g_t \le \bar{g} \quad (t \in \mathcal{T}_0), \quad (4)$$

$$\underline{l} \le l_t \le \bar{l} \quad (t \in \mathcal{T}_0), \quad (5)$$

$$0 \le C_0 + \sum_{\tau=1}^{t}(\kappa h_\tau^{l*} l_\tau - h_\tau^{g*} g_\tau) \le \bar{C} \quad (t \in \mathcal{T}_0), \quad (6)$$

$$h_t^{g*} = \begin{cases} 0, & \text{if } C_0 + \sum_{\tau=1}^{t}(\kappa h_\tau^{l*} l_\tau - h_\tau^{g*} g_\tau) = 0 \wedge \hat{h}_t^g = 1 \\ \hat{h}_t^g, & \text{otherwise,} \end{cases} \quad (t \in \mathcal{T}_0), \quad (7)$$

$$h_t^{l*} = \begin{cases} 0, & \text{if } C_0 + \sum_{\tau=1}^{t}(\kappa h_\tau^{l*} l_\tau - h_\tau^{g*} g_\tau) = \bar{C} \wedge \hat{h}_t^l = 1 \\ \hat{h}_t^l, & \text{otherwise,} \end{cases} \quad (t \in \mathcal{T}_0), \quad (8)$$

where $\underline{g}$ and $\bar{g}$ are the lower/upper limits of the PSHG generation during pump-down, respectively, $\underline{l}$ and $\bar{l}$ are the lower/upper limits of the PSHG load for pump-up, respectively, C_0 is the initial water level of the upper reservoir for the control period $\mathcal{T}_0$, $\bar{C}$ represents the maximum capacity of the reservoir, and κ is the operating efficiency of the PSHG. Under the control policy shown in Eqs. (3) – (8), the BG operation scheduling problem for the day-ahead market can be formulated as a problem of determining the decision variable set $\{(b_t, \boldsymbol{h}_t); t \in \mathcal{T}_0\}$.

III. SCHEDULING OF BG OPERATION

A. OVERVIEW OF BG SCHEDULING

To realize sustainable energy management while taking into account the impacts of uncertainty, a framework such as receding horizon control [38] is effective, in which plans derived during the operational scheduling process are modified and updated based on the latest prediction results to appropriately utilize energy buffers with limited capacity [39], [40]. This study follows this concept and assumes the timeline shown in Fig. 1. The framework uses forecast results for 96 timeslots representing the next 48 hours using information available at 08:00 AM of the previous day; develops a BG operational plan for period $\mathcal{T}_+$, which includes the target period $\mathcal{T}_0$; and submits the operational plan for $\mathcal{T}_0$ to the market. The water level of the PSHG at the beginning of the target planning period, C_0, is estimated using the wind power output prediction results over $\mathcal{T}_-$. In summary, the following input-output relationship represents our proposed BG scheduling task:

$$\{\{\Pr(W_t); t \in \mathcal{T}\}, \{(\hat{c}_t, \hat{i}_t); t \in \mathcal{T}_+\}\} \to \{(b_t, \boldsymbol{h}_t); t \in \mathcal{T}_0\}, \quad (9)$$

where $\Pr(W_t)$ represents the probability density function of the predicted wind power output, and $\hat{c}_t$ and $\hat{i}_t$ are the estimated unit values of the day-ahead market price and imbalance charge, respectively.[2] This study mainly focuses on the effective use of uncertainty information in wind power outputs in BG planning under the given forecast results of the market price and imbalance charge. Figure 2 shows the schematic of the proposed BG operational planning framework.

In the remaining subsections, we first introduce the PDP framework to represent heterogeneity in WF output uncertainty. Then, the target BG operational planning problem is described; in particular, we introduce the objective function involving numerical integration operations and the PSHG capacity chance-constraints representation due to the time dependence of the uncertainty in the WF output to identify challenges in the optimization process. Then, to represent these chance constraints explicitly, a copula modeling approach is introduced for generating plausible WF output scenarios. Finally, we introduce surrogate functions representing the upper/lower bounds of the original objective function and propose an iterative procedure to achieve optimization while efficiently inducing these bounds to be tight.

B. PDP OF WF OUTPUT

We focus on the probability density prediction (PDP) scheme introduced in [31] for the target WF output, which is based on the methodology called NGBoost [41]. In this scheme, the probability density of the WF output for each 30-min

[2]Note that the JEPX day-ahead market is traded via the *single-price auction* scheme, so the sensitivity of profitability to bid prices for BG operators is low.

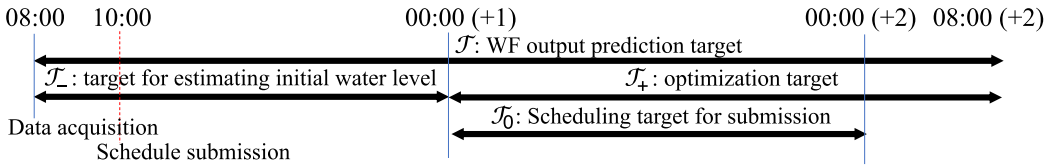


FIGURE 1. Scheduling timeline of BG operation.

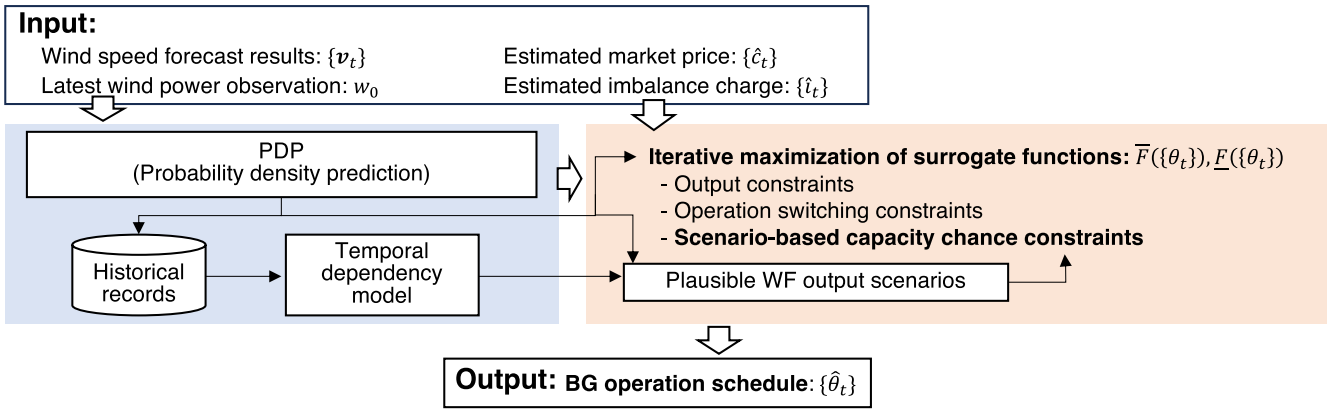


FIGURE 2. Proposed BG operational scheduling scheme.

slot in the target period, $\mathcal{T}(\supset \mathcal{T}_0)$, is derived based on the numerical weather prediction (NWP) result [42] of the corresponding timeslot and latest WF power generation observation at the time of prediction; the NWP result used in this scheme consists of horizontal wind speed information (two-dimensional information composed of the zonal and meridional velocities) at an altitude of 60 m on 15 km mesh points within 40 km of the target site provided by an ensemble weather forecast (EWF) [27], [30] with 11 members. The schematic of the information flow in the PDP scheme used in this study is shown in Figure 3.

In this PDP scheme, we focus on the statistical behavior of the per-unit WF power generation, i.e.,

$$y_t = \frac{w_t - \min_{t'}(w_{t'})}{\max_{t'}(w_{t'}) - \min_{t'}(w_{t'})}. \tag{10}$$

Let $\boldsymbol{v}_t$ be the vectorized wind speed information for timeslot $t \in \mathcal{T}$ in the latest EWF results accessible at the time of prediction. Our PDP scheme derives the conditional probability density of the WF output in the target timeslot, $t \in \mathcal{T}$, using information $\boldsymbol{v}_t$ and latest information y_0 as follows,

$$\Pr(y_t \mid \boldsymbol{v}_t, y_0) = p_t(y_t; \boldsymbol{\phi}_t(\boldsymbol{v}_t, y_0)). \tag{11}$$

In this study, we focus on the beta distribution as a parametric model for such a description[3] and consider deriving parameter $\boldsymbol{\phi}_t$ to represent this conditional distribution using gradient boosting trees [36] based on the natural gradient concept [45] whose inputs consist of $\boldsymbol{v}_t$ and y_0, which can be observed at the time of prediction. Algorithm 1 shows the learning procedure of the PDP model in terms of NGBoost. The algorithm derives a set of component models, $\{\chi_t^{(k)}\}$, and their coefficients, $\{\rho_t^{(k)}\}$, to represent the parameter of the conditional probability density function, i.e. Eq. (11); the derived parameter is represented as

$$\hat{\boldsymbol{\phi}}_t(\boldsymbol{v}_t, y_0) = \hat{\boldsymbol{\phi}}_t^{(0)} - \eta \sum_{k=1}^{K} \rho_t^{(k)} \chi_t^{(k)}(\boldsymbol{v}_t, y_0), \tag{12}$$

where η is a learning rate (e.g., $\eta = 0.01$). In Algorithm 1, the learning process is carried out by focusing on the log-likelihood of the probability density, i.e., shown as $l(.)$, and the Fisher Information shown as $I(.)$ under the current model. The model is learned based on the historical dataset $\mathcal{D}_t = \{(\boldsymbol{v}_{d,t}, y_{d,0}), y_{d,t}; d \in \mathcal{L}\}$, where $\boldsymbol{v}_{d,t}$ and $y_{d,t}$ indicate information of $t \in \mathcal{T}$ in day d in the learning period $\mathcal{L}$, and $y_{d,0}$ indicates the latest information of WF power generation at the prediction timing in day d in the learning period $\mathcal{L}$. Note that by adopting the inverse operation of Eq. (10), we can easily represent the conditional probability of the original wind power generation, i.e., $\Pr(w_t \mid \boldsymbol{v}_t, w_0) = p_t(w_t; \phi_t(\boldsymbol{v}_t, w_0))$; we write simply $\Pr(w_t)$ when no confusion can arise. We stress that this approach has the ability to represent *heterogeneous* uncertainties that vary with the situation for each timeslot.

[3] Since it has been reported that the PDP results under the assumption of the beta distribution better represent the stochastic behavior of real-world WF outputs [31] in terms of the *rank histograms* [43] and the *continuous ranked probability score* (CRPS) [44] than, for example, those under the assumption of the Gaussian distribution or those utilizing a nonparametric distribution derived with the bootstrap sampling technique, we adopted the beta distribution here.

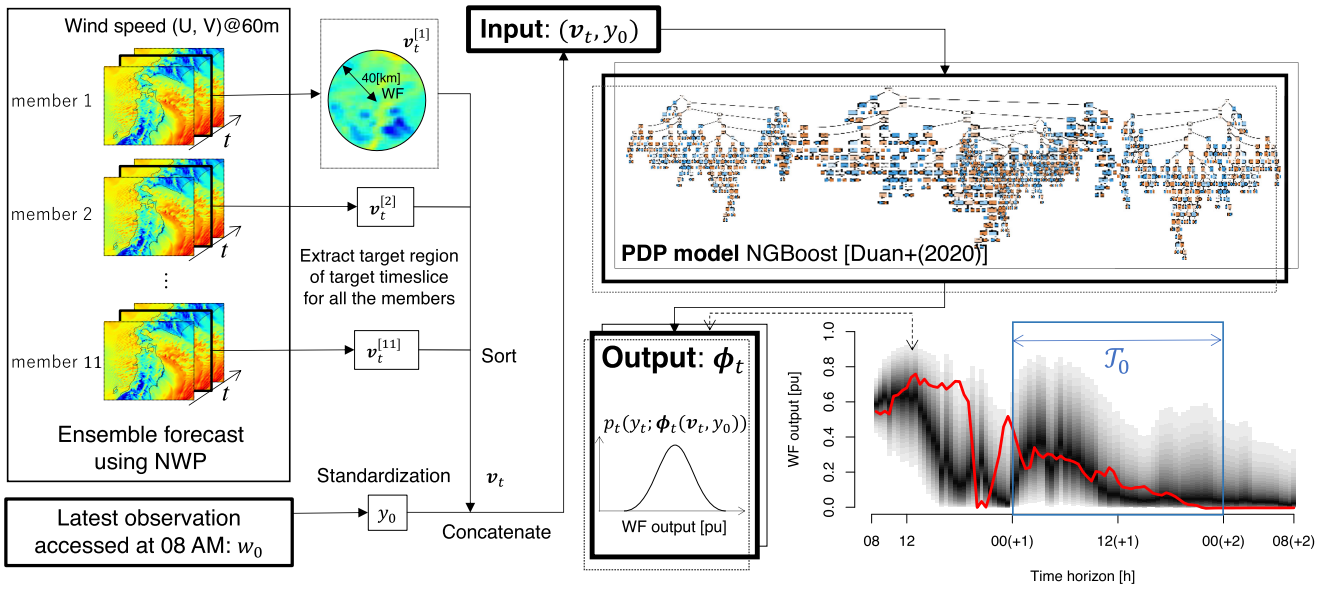


FIGURE 3. Schematics of the information flow of the PDP scheme used in this study (see [31] for detail).

Algorithm 1 Learning Probabilistic Prediction Models

Input: $\mathcal{D}_t = \{(\boldsymbol{v}_{d,t}, y_{d,0}), y_{d,t}\}_{d\in\mathcal{L}}$.

$\boldsymbol{\phi}_t^{(0)} \leftarrow \underset{\boldsymbol{\phi}}{\operatorname{argmin}} \sum_{d\in\mathcal{L}} l(\boldsymbol{\phi}, y_{d,t})$.

$\boldsymbol{\phi}_d^{(0)} \leftarrow \boldsymbol{\phi}_t^{(0)} \quad (\forall d \in \mathcal{L})$.

for $k = 1, \ldots, K$ **do**

  $\alpha_d^{(k)} \leftarrow I(\boldsymbol{\phi}_d^{(k-1)})^{-1} \frac{\partial}{\partial \boldsymbol{\phi}} l(\boldsymbol{\phi}_d^{(k-1)}, y_{d,t}) \quad (\forall d \in \mathcal{L})$.

  $\chi_t^{(k)} \leftarrow \mathrm{fit}(\{(\boldsymbol{v}_{d,t}, y_{d,0}), \alpha_d^{(k)}\}_{d\in\mathcal{L}})$.

  $\rho_t^{(k)} \leftarrow \underset{\rho}{\operatorname{argmin}} \sum_{d\in\mathcal{L}} l(\boldsymbol{\phi}_d^{(k-1)} - \rho\chi_t^{(k)}(\boldsymbol{v}_{d,t}, y_{d,0}), y_{d,t})$.

  $\boldsymbol{\phi}_d^{(k)} \leftarrow \boldsymbol{\phi}_d^{(k-1)} - \eta(\rho_t^{(k)}\chi_t^{(k)}(\boldsymbol{v}_{d,t}, y_{d,0})) \quad (\forall d \in \mathcal{L})$.

end for

Output: $\{\boldsymbol{\phi}_t^{(0)}, \{\rho_t^{(k)}, \chi_t^{(k)}\}_{k=1}^K\}$.

C. OPTIMIZATION OF BG SCHEDULING

Here, we formally introduce the target BG operation planning problem based on maximizing the expected revenue. We write $\theta_t = (b_t, \boldsymbol{h}_t, g_t, l_t)$ for the decision variable set for the planning at timeslot t; the BG scheduling problem is then defined as follows:

$$\begin{aligned}\{\hat{\theta}_t\} &= \underset{\{\theta_t\}}{\operatorname{argmax}}\, F(\{\theta_t\}) \\ &= \underset{\{\theta_t\}}{\operatorname{argmax}} \sum_{t\in\mathcal{T}_+} \mathbb{E}\left[\hat{r}_t(W_t \mid \theta_t)\right] \\ &= \underset{\{\theta_t\}}{\operatorname{argmax}} \sum_{t\in\mathcal{T}_+} \int \Pr(w_t)\hat{r}_t(w_t \mid \theta_t)dw_t, \end{aligned} \tag{13}$$

where $\hat{r}_t(w_t \mid \theta_t)$ is the estimated revenue, and

$$\hat{r}_t(w_t \mid \hat{\theta}_t) = \hat{c}_t b_t - \hat{i}_t \left[b_t - (w_t + h_t^g \bar{g} - h_t^l l)\right]_+ . \tag{14}$$

The first term of Eq. (14) represents the revenue or expenditure in the day-ahead market, whereas the second term represents the payment for the imbalance charge. For the latter, when the deviation from the scheduled value of the BG output can be compensated for by controlling the speed of PSHG generation/loading or curtailing the output of the WF, no imbalance occurs, and the charge becomes zero.

The constraints of the BG scheduling problem, Eq. (13), introduced in this study are given as follows.

Output constraints:

$$\underline{g} \le g_t \le \bar{g} \quad (t \in \mathcal{T}_+) \tag{15}$$

$$\underline{l} \le l_t \le \bar{l} \quad (t \in \mathcal{T}_+) \tag{16}$$

Operation switching constraints:

$$h_t^g + h_t^l \le 1 \quad (t \in \mathcal{T}_+) \tag{17}$$

$$h_{t-1}^l + h_t^g \le 1 \quad (t \in \mathcal{T}_+) \tag{18}$$

$$h_{t-1}^g + h_t^l \le 1 \quad (t \in \mathcal{T}_+) \tag{19}$$

$$\sum_{t=1}^{|\mathcal{T}|} \{|h_t^g - h_{t-1}^g| + |h_t^l - h_{t-1}^l|\} \le B \tag{20}$$

Here, B is the limit of the operation switching frequency for the PSHG.

Capacity chance constraint:

$$\Pr\left(0 \le \hat{C}_0 + \sum_{\tau=1}^{t}\left(\kappa h_\tau^l l_\tau - h_\tau^g g_\tau\right) \le \bar{C}\right) \ge \zeta \quad (t \in \mathcal{T}_+) \tag{21}$$

here, $\zeta \in (0, 1)$ is a threshold parameter, and $\hat{C}_0$ is the estimation result of the initial storage capacity at the start time of the scheduling period. In this study, the initial storage

capacity is estimated as follows:

$$\hat{C}_0 = C_0' + \sum_{t \in \mathcal{T}_-} (\kappa \hat{h}_t^{l'} \hat{l}_t^* - \hat{h}_t^{g'} \hat{g}_t^*), \tag{22}$$

where C_0' is the current capacity observable at the scheduling time, $\hat{h}_t^{g'}$ and $\hat{h}_t^{l'}$ are plausible operation states determined under the schedule that was submitted the day before, and $\hat{g}_t^*$ and $\hat{l}_t^*$ suggest the plausible control results. These plausible parameters are given by deriving the output of the control phase for period $\mathcal{T}_-$ using the expected WF output derived from the PDP in a manner similar to that shown in Eqs. (3)–(8).

Note that while the variables of substantial importance in our BG scheduling task are $(b_t, \boldsymbol{h}_t)$, we use $\theta_t = (b_t, \boldsymbol{h}_t, g_t, l_t)$ as the decision variable set for the optimization in evaluating the feasibility of the solution for this capacity constraint. As described in Section II, our framework assumes that $\{(b_t, \boldsymbol{h}_t); t \in \mathcal{T}_0\}$, which corresponds to the subset of the optimizer, is adopted as the scheduling result and that the real-time output control is conducted according to $\{(g_t^*, l_t^*)\}$, which is derived in Eq. (3).

The technical difficulties in implementing this framework are the integral calculation for evaluating the expected value of the objective function $F(\{\theta_t\})$ appearing in Eq. (13) in the optimization process and the explicit evaluation of the chance constraint shown in Eq. (21) by considering the uncertainty relationships among timeslots, i.e., *time-dependence*. To address these difficulties, we introduce the representation of stochastic scenario-based capacity chance constraints using a copula function [46] and then propose a computationally efficient iterative optimization approach by introducing surrogate functions to evaluate the upper/lower bounds of the objective function $F(\{\theta_t\})$.

D. PSHG CAPACITY CONSTRAINTS BASED ON WF OUTPUT SCENARIOS

The advantage of the BG operation scheme is that deviations from the wind generation forecast can be mitigated within the controllability of the PSHG. However, the temporal behavior of the realized wind power generation relative to the forecast result generally has a time-dependence; e.g., when a measurement is over/under-performing relative to the PDP result in a timeslot t, the measurement tends to behave similarly in $t+1$. Meanwhile, the measurement rarely continues to deviate in the same direction from the PDP result for the entire timeslot $\mathcal{T}_0$. Therefore, to elaborate on the capacity constraint affected by the accumulation of output results over the target period, as shown in Eq. (21), the time-dependence of the relative position of WF output results to the PDP results must be appropriately considered.

In this study, the following steps are introduced to allow the evaluation of the capacity constraint in an explicit form while neatly incorporating such time-dependence.

1) Statistically model the temporal dependencies of the behavior of the realized wind power generation on the PDP results.
2) Generate multiple scenarios based on the statistical model.
3) Represent the stochastic chance constraints on the capacity using the generated scenarios.

In Step 1, we represent the joint probability density function of wind power generation in the target period, $\{w_t; t \in \mathcal{T}_+\}$, as follows:

$$\Pr(w_1, \ldots, w_{|\mathcal{T}_+|}) = \psi(\xi_1, \ldots, \xi_{|\mathcal{T}_+|}) \prod_{t \in \mathcal{T}_+} \Pr(w_t), \tag{23}$$

where ξ_t represents the quantile from the viewpoint of the marginal cumulative distribution of the PDP result, i.e.,

$$\xi_t = \Pr(W_t < w_t) \quad (t \in \mathcal{T}_+), \tag{24}$$

and $\psi(.)$ is a copula density function [46]. In this study, we focus on the copula density function modeled via a regular vine (R-vine) copula [35], flexibly representing the complex dependence structure among multiple timeslots with a combination of various types of paired copulas (see Appendix A for a brief description of the R-vine). This copula density function is trained offline before the scheduling phase using historical records of past PDP results and actual WF output behaviors.

In Step 2, multiple WF output scenarios are generated via inverse transformation sampling [47] using the PDP results available at the scheduling phase, $\{\Pr(w_t); t \in \mathcal{T}_+\}$, and copula density function, $\psi(\xi_1, \ldots, \xi_{|\mathcal{T}_+|})$, which is estimated in Step 1. In this step, J scenarios are generated as follows:

$$(\hat{w}_1^{(j)}, \ldots, \hat{w}_{|\mathcal{T}_+|}^{(j)}) \sim \Pr(W_1, \ldots, W_{|\mathcal{T}_+|}) \quad (j = 1, .., J). \tag{25}$$

Finally, in Step 3, we evaluate the J scenarios in terms of the log-likelihood based on the model,

$$\log \Pr(\hat{w}_1^{(j)}, \ldots, \hat{w}_{|\mathcal{T}_+|}^{(j)}) \quad (j = 1, \ldots, J), \tag{26}$$

extract a scenario set $\mathcal{S}$ consisting of S scenarios with the highest log-likelihood among J generated scenarios, and introduce the following scenario-based capacity constraints that consider the feasibility of these scenarios at a minimum.

Scenario-based capacity chance constraints:

$$0 \le \max_{t \in \mathcal{T}_+, s \in \mathcal{S}} C_t^{(s)} \le \bar{C}, \tag{27}$$

$$0 \le \min_{t \in \mathcal{T}_+, s \in \mathcal{S}} C_t^{(s)} \le \bar{C}, \tag{28}$$

where

$$C_t^{(s)} = C_0 + \sum_{\tau=1}^{t} \Big[-\kappa h_t^l \min\{\max\{b_\tau - \hat{w}_\tau^{(s)}, -\bar{l}\}, -\underline{l}\} - h_\tau^g \min\{\max\{b_\tau - \hat{w}_\tau^{(s)}, \underline{g}\}, \bar{g}\} \Big] \quad (t \in \mathcal{T}_+, s \in \mathcal{S}), \tag{29}$$

and $C_t^{(s)}$ are the PSHG water level in scenarios $s \in \mathcal{S}$.

Constraints shown in Eqs. (27) and (28) ensure that the maximum and minimum values of the water level transition among all S scenarios in each timeslot are within the feasible

storage capacity range. Note that these alternative constraints imply that the ratio of S to J acts like the role of the parameter ζ in Eq. (21) and enforce not to violate the capacity constraints for at least S scenarios among J. Instead of the capacity chance constraint shown in Eq. (21), we use the constraints shown in Eqs. (27) and (28), which can be evaluated explicitly.

E. ITERATIVE OPTIMIZATION USING SURROGATE FUNCTIONS

The expected value optimization problem introduced in Eq. (13) involves integral calculation, which makes direct numerical optimization computationally expensive. To address this computational problem, we introduce surrogate functions that express upper/lower bounds via discrete bounding approximation such that the objective function $F(\{\theta_t\})$ in Eq. (13) can be optimized efficiently and in practical time (i.e., 1–2 hours) by iteratively tightening the difference between these bounds while narrowing the solution search space.

First, we discretize the domain of the wind power output at timeslot t by dividing it into N_t parts. Then, we derive the upper/lower bounds of the expected value in the partitioned domain $[\underline{w}_t^n, \bar{w}_t^n]$ for all $n \in \mathcal{N}_t (= \{1, \ldots, N_t\})$, where $\underline{w}_t^n$ and $\bar{w}_t^n$ are the quantiles and have the following properties,

$$\Pr\left(W_t \leq \underline{w}_t^n\right) = \frac{n-1}{N_t}, \tag{30}$$

$$\Pr\left(W_t \leq \bar{w}_t^n\right) = \frac{n}{N_t}. \tag{31}$$

We introduce a surrogate function $\bar{F}(\{\theta_t\}; \mathcal{N}_t) \geq F(\{\theta_t\})$ representing the upper bound given the partitioning set $\{\mathcal{N}_t = \{1, \ldots, N_t\}; t \in \mathcal{T}_+\}$ via discrete bounding approximation (see Appendix B for details) as follows:

$$\begin{aligned}\bar{F}(\{\theta_t\}; \{\mathcal{N}_t\}) &= \sum_{t \in \mathcal{T}_+} \bar{F}_t(\theta_t; \mathcal{N}_t) \\ &= \sum_{t \in \mathcal{T}_+} \frac{1}{N_t} \sum_{n \in \mathcal{N}_t} \hat{r}(\tilde{w}_t^n \mid \theta_t),\end{aligned} \tag{32}$$

where

$$\tilde{w}_t^n = N_t \int_{\underline{w}_t^n}^{\bar{w}_t^n} w_t \Pr(w_t) dw_t. \tag{33}$$

Similarly, the following surrogate function is introduced to represent the lower bound (see Appendix B for details) with the property $\underline{F}(\{\theta_t\}; \{\mathcal{N}_t\}) \leq F(\{\theta_t\})$:

$$\begin{aligned}&\underline{F}(\{\theta_t\}; \{\mathcal{N}_t\}) \\ &= \sum_{t \in \mathcal{T}_+} \underline{F}_t(\theta_t; \mathcal{N}_t) \\ &= \sum_{t \in \mathcal{T}_+} \sum_{n \in \mathcal{N}_t} \left\{ \frac{\bar{w}_t^n - \tilde{w}_t^n}{\bar{w}_t^n - \underline{w}_t^n} \hat{r}(\underline{w}_t^n | \theta_t) + \frac{\tilde{w}_t^n - \underline{w}_t^n}{\bar{w}_t^n - \underline{w}_t^n} \hat{r}(\bar{w}_t^n | \theta_t) \right\}.\end{aligned} \tag{34}$$

Note that $\underline{w}_t^n$, $\bar{w}_t^n$, and $\tilde{w}_t^n$ are explicitly calculable under the PDP results. Therefore, the maximization problem of each surrogate function with constraints Eqs. (15)–(20) and (27), (28) takes the form of the mixed integer linear program (MILP); thus, given an appropriate partitioning set $\{\mathcal{N}_t\}$ that achieves sufficiently tight upper/lower bounds and using a modern numerical solver, e.g., Gurobi, we can derive the well-approximated maximizer of the original objective function, $F(\{\theta_t\})$, which is challenging to optimize directly.

To achieve this optimization within acceptable computation time in practical situations while finding a suitable partitioning set that achieves a sufficiently tight discrete approximation, we propose the iterative optimization procedure shown in Algorithm 2. The algorithm first derives the optimal solution sets for the surrogate functions[4] introduced in Eqs. (32) and (34) (lines 9 and 10 in Algorithm 2) under a rough partitioning set $\{\mathcal{N}_t\}$ given at the beginning (line 1); these maximizers can be derived independently by parallel computation. Suppose the optimization results for two surrogate functions show that the operating condition of the PSHG, $\boldsymbol{h}_t$, at a certain timeslot t is consistent and the difference between the upper/lower bounds is sufficiently tight; in that case, this operation condition is fixed, adopted for this timing, and excluded from the optimization in the subsequent procedure (lines 11–17). Moreover, for timeslots with large differences between the upper/lower bounds, the resolutions of the discretization are improved (line 18), and the surrogate functions are optimized again using a redesigned partitioning set $\{\mathcal{N}_t\}$. These processes are performed iteratively. After the convergence condition is satisfied, the upper and lower boundary solutions are substituted into the original objective function, and the solution with the larger objective function value is adopted as the final solution (line 20).

IV. NUMERICAL EXPERIMENTS

A. EXPERIMENTAL SETUP

Annual simulations were conducted using a dataset collected at a real-world WF located in eastern Japan from January 1, 2016, to January 31, 2018, and from historical datasets of JEPX day-ahead market prices[5] and imbalance charges[6] to evaluate the effectiveness of the proposed scheme. To evaluate the impact of the BG scheme on achievable revenue, the following five approaches were compared:

- *Independent operation*: The WF and PSHG are scheduled and controlled independently; the PSHG does not compensate for deviations between the scheduled

[4] Note that although $\{g_t\}$ and $\{l_t\}$ are not explicitly described as arguments of the surrogate functions in Algorithm 2 to simplify the notation, they are also optimization targets in the procedure.

[5] "Trading information," [Online] Available: https://www.jepx.jp/electricpower/market-data/spot/, accessed: 2023-08-02.

[6] "Unit price of imbalance charge," [Online] Available: https://www.tepco.co.jp/pg/consignment/retailservice/imbalance, accessed: 2023-08-02.

Algorithm 2 Iterative Optimization Procedure Based on Surrogate Functions

Input: $\hat{c}_t, \hat{i}_t, \Pr(w_t)\ \forall t$, and $\hat{C}_0$.

$N_t \leftarrow N^0, \quad \Delta_t^{\mathrm{rev}} \leftarrow \varepsilon^{\mathrm{rev}} + 1, \quad \Delta_t^{\mathrm{bg}} \leftarrow \varepsilon^{\mathrm{bg}} + 1 \quad (\forall t)$.
$\hat{\boldsymbol{h}}_t \leftarrow \emptyset\ (\forall t), \mathcal{T}^* \leftarrow \emptyset$.
while $\{t; (\Delta_t^{\mathrm{rev}} > \varepsilon^{\mathrm{rev}} \vee \Delta_t^{\mathrm{bg}} > \varepsilon^{\mathrm{bg}}) \wedge N_t \leq \bar{N}\} \neq \emptyset$ **do**
  for $t \in \{t; (\Delta_t^{\mathrm{rev}} > \varepsilon^{\mathrm{rev}} \vee \Delta_t^{\mathrm{bg}} > \varepsilon^{\mathrm{bg}}) \wedge N_t \leq \bar{N}\}$ **do**
    for $n \in \mathcal{N}_t (= \{1, \ldots, N_t\})$ **do**
      Calculate Eqs. (30), (31), and (33).
    end for
  end for
  $\{\{\bar{b}_t\}, \{\bar{\boldsymbol{h}}_t\}\} \leftarrow \underset{\{b_t\},\{\boldsymbol{h}_t\}}{\mathrm{argmax}} \sum_{t \in \mathcal{T}_+ \setminus \mathcal{T}^*} \bar{F}_t(b_t, \boldsymbol{h}_t; \mathcal{N}_t) + \sum_{t \in \mathcal{T}^*} \bar{F}_t(b_t, \hat{\boldsymbol{h}}_t; \mathcal{N}_t)$, subject to Eqs. (15)–(20), (27), and (28).
  $\{\{\underline{b}_t\}, \{\underline{\boldsymbol{h}}_t\}\} \leftarrow \underset{\{b_t\},\{\boldsymbol{h}_t\}}{\mathrm{argmax}} \sum_{t \in \mathcal{T}_+ \setminus \mathcal{T}^*} \underline{F}_t(b_t, \boldsymbol{h}_t; \mathcal{N}_t) + \sum_{t \in \mathcal{T}^*} \underline{F}_t(b_t, \hat{\boldsymbol{h}}_t; \mathcal{N}_t)$, subject to Eqs. (15)–(20), (27), and (28).
  $\bar{\theta}_t \leftarrow \begin{cases} \{\hat{b}_t, \hat{\boldsymbol{h}}_t\} & \text{if } t \in \mathcal{T}^* \\ \{\bar{b}_t, \bar{\boldsymbol{h}}_t\} & \text{otherwise} \end{cases}, \quad \underline{\theta}_t \leftarrow \begin{cases} \{\hat{b}_t, \hat{\boldsymbol{h}}_t\} & \text{if } t \in \mathcal{T}^* \\ \{\underline{b}_t, \underline{\boldsymbol{h}}_t\} & \text{otherwise} \end{cases}$.
  $\Delta_t^{\mathrm{rev}} \leftarrow \bar{F}_t(\bar{\theta}_t; \mathcal{N}_t) - \underline{F}_t(\underline{\theta}_t; \mathcal{N}_t)\ (\forall t), \quad \Delta_t^{\mathrm{bg}} \leftarrow |\bar{b}_t - \underline{b}_t|\ (\forall t)$.
  if $\{t; \bar{\boldsymbol{h}}_t = \underline{\boldsymbol{h}}_t\} = \mathcal{T}_+$ **then**
    $\mathcal{T}^* \leftarrow \mathcal{T}_+, \quad \hat{\boldsymbol{h}}_t \leftarrow \bar{\boldsymbol{h}}_t (= \underline{\boldsymbol{h}}_t)\ (t \in \mathcal{T}_+)$.
  else
    $\mathcal{T}^* \leftarrow \{t; \bar{\boldsymbol{h}}_t = \underline{\boldsymbol{h}}_t \wedge \Delta_t^{\mathrm{rev}} < \mathrm{median}_{t' \in \mathcal{T}_+}\{\Delta_{t'}^{\mathrm{rev}}\}\}$, and $\hat{\boldsymbol{h}}_t \leftarrow \bar{\boldsymbol{h}}_t\ (t \in \mathcal{T}^*)$.
  end if
  $N_t \leftarrow 2 \times N_t \quad (t \in \{t; (\Delta_t^{\mathrm{rev}} > \varepsilon^{\mathrm{rev}} \vee \Delta_t^{\mathrm{bg}} > \varepsilon^{\mathrm{bg}}) \wedge N_t \leq \bar{N}\})$.
end while
$\hat{\theta} \leftarrow \begin{cases} \{\bar{\theta}_t\} & \text{if } \sum_{t \in \mathcal{T}_+} F_t(\bar{\theta}_t, \mathcal{N}_t) \geq \sum_{t \in \mathcal{T}_+} F_t(\underline{\theta}_t, \mathcal{N}_t) \\ \{\underline{\theta}_t\} & \text{otherwise.} \end{cases}$

Output: $\hat{\theta}$.

and actual WF outputs. The WF output schedule was given as the expected values derived from the forecast results, and the PSHG output schedule was determined to maximize its revenue in the day-ahead market without considering the uncertainty in the WF output. This approach follows the traditional operation from the literature [17], [18], [19].

- *Naive BG*: A naive BG scheme without active cooperation, in which the operation schedules of the WF and PSHG are determined separately, and deviations in the WF output are compensated by the PSHG in the control phase. This approach follows the benchmarking approach introduced in the literature [2].
- *Coordinated wind-pumped storage hydropower (CWPSH) operation* [1]: A state-of-the-art cooperation approach proposed in [1]. The day-ahead operation schedules of the WF and PSHG are planned in a coordinated manner; the impact of uncertainty in wind power output on PSHG capacity is considered in the constraints during the scheduling phase, but the uncertainty in WF forecast results are represented by the common single empirical probability density function (EPDF) of the historical prediction errors around the deterministic prediction results[7] under the assumption of homogeneity; the time-dependence of the uncertainty in WF forecast results is also not considered.
- *Hydro-wind joint bidding (HWJB) operation* [2]: Another state-of-the-art approach of the hydro-wind joint bidding proposed in [2]. A cooperation scheme in which the operation schedules of the WF and PSHG are determined by considering the uncertainty in the WF output, and deviations in the WF output are compensated by the PSHG in the control phase. The time-dependent uncertainty in the WF output is represented by the PDP results; thus, the heterogeneous uncertainties are represented in the scheduling phase. The time-dependence of the uncertainty in WF forecast results is not considered.
- *Proposed BG operation*: The proposed BG scheme aiming to maximize the expected revenue. The cooperation scheme considers the time-dependence in the

[7] In this experiment, we used the expected value of the distribution derived from the prediction mechanism used in Section III-B as the deterministic prediction result. It has been reported in [31] that the predictions derived in this way are comparable in accuracy to deterministic prediction methods based on the state-of-the-art mechanism.

TABLE 2. BG operational parameters used in the experiment.

Description	Variable	Value
WF	Rated power output	340 MW
PSHG	$\bar{g}, \underline{g}$	150, 75 MW
	$\bar{l}, \underline{l}$	150, 105 MW
	$\bar{C}$	2010 MWh
	C_0 (for the first day)	0 MWh
	B, κ	4, 0.7
Optimization parameters	$\varepsilon^{\mathrm{rev}}, \varepsilon^{\mathrm{bg}}$	500, 10
	$N^0, \bar{N}$	5, 320

heterogeneous uncertainties in WF forecast results for the BG scheduling.

The unit values of the day-ahead market price $\hat{c}_t$ and imbalance charge $\hat{i}_t$ used in the scheduling phase were naively estimated using the prices of the most recent day, distinguishing weekdays and holidays, as follows:

$$\hat{c}_t = c'_t \quad (t \in \mathcal{T}_0,) \tag{35}$$

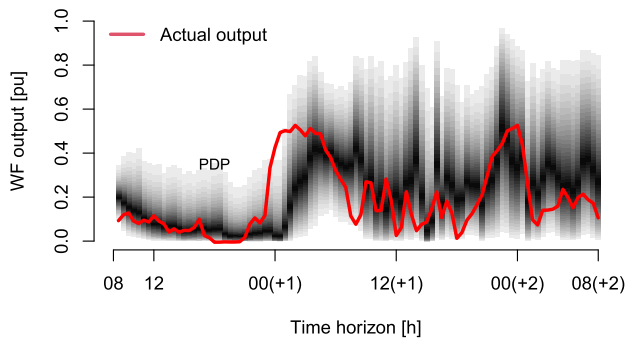
$$\hat{i}_t = \begin{cases} i'_t & \text{if } (i'_t > c'_t) \\ c'_t + 10 & \text{otherwise} \end{cases} \quad (t \in \mathcal{T}_0), \tag{36}$$

where c'_t and i'_t are the day-ahead market price and imbalance charge in yen/MWh on the day referred to, respectively. Table 2 presents the specifications of the WF and PSHG that constitute the BG and the parameters used in the operational scheduling optimization. This experiment focused on the annual data subset from February 1, 2017, to January 31, 2018, as the evaluation period. The prediction model of the WF was updated once a month using the data from January 1, 2016, through the previous month of the operational target period. In the proposed scheme, scenarios for the PSHG capacity constraints were generated by modeling copulas based on a historical dataset of PDP results and actual outputs of the WF collected one month before the operation date. The annual operational performance of each approach was evaluated by simulating the continuous daily BG operation succeeding the previous day's water level based on the data collected and predicted for this evaluation period. All experiments were performed using Gurobi 9.5.1 as the optimization solver and were evaluated in a parallel computing environment with eight threads using an Intel Xeon Platinum 8258 processor and 512GB of memory.

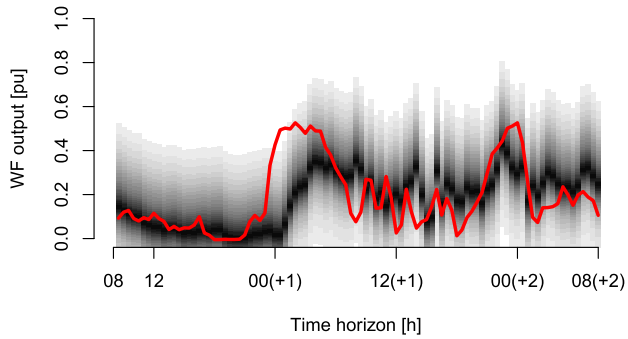
B. EVALUATION OF PDP RESULTS

Firstly, we show the characteristics of the PDP results for the WF output[8] in terms of the heterogeneity of the uncertainty represented in the prediction results and their time-dependence. Figure 4 shows examples of the prediction intervals for various percentiles under the resulting probability densities from 30 minutes to 48 hours ahead. Figure 4(a) suggests that the PDP framework introduced in

[8]While the further detailed discussion of the PDP performance is beyond the scope of this paper, a quantitative evaluation of the PDP scheme implemented in this study was reported in [31]; we refer readers to it for evaluation results under conditions comparable to those of this experiment.



(a) Representative PDP result.

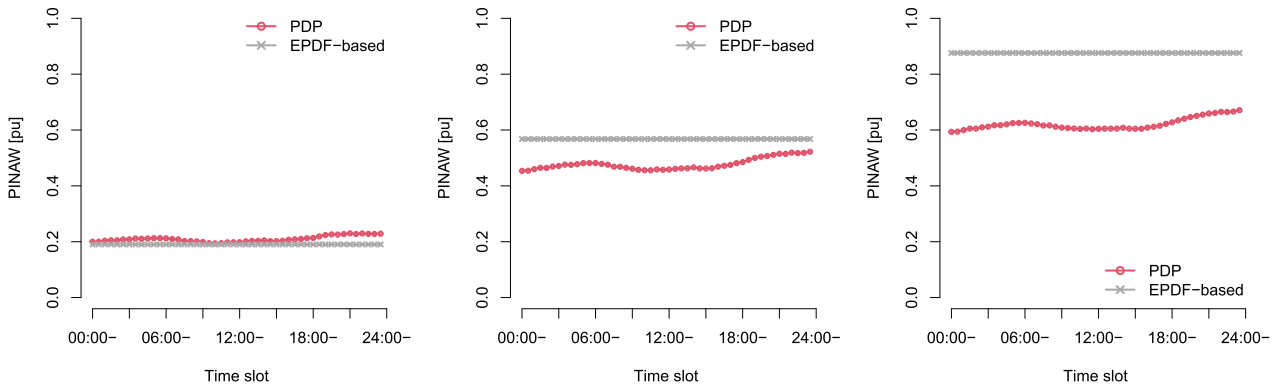


(b) Uncertainty represented by common single EPDF.

FIGURE 4. Representative results of PDP used in our approach and uncertainty.

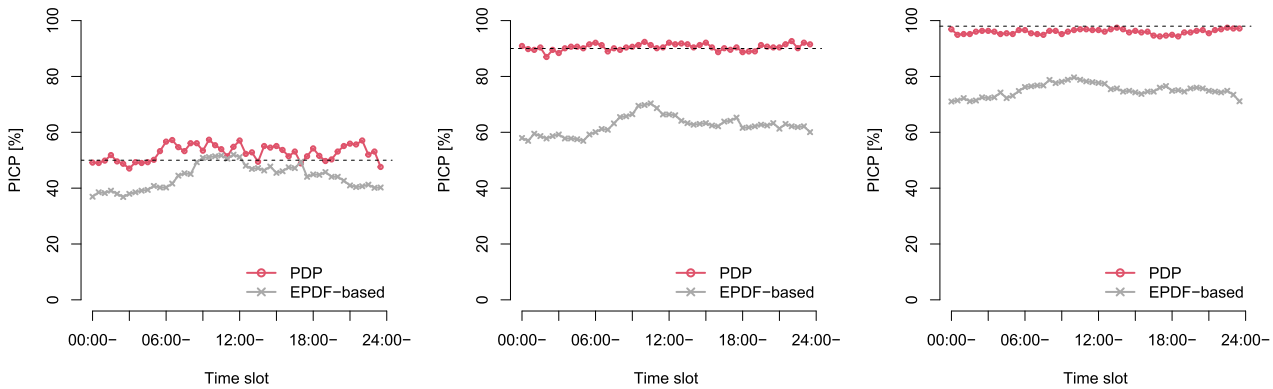
Section III-B can represent heterogeneity in the uncertainty of WF output, such that the width and bias of the derived prediction intervals vary depending on the situation. For example, the results of the proposed PDP confidently suggest low WF outputs around 08:00–23:00 on the first day by peaky probability distributions, while expressing low confidence in the WF outputs around 10:00–13:00 on the next day with broad, less peaky probability distributions. On the other hand, Fig. 4(b) shows the forecast results represented by the common single EPDF constructed with historical error trends of the deterministic prediction results used in CWPSH; this also represents the uncertainty in the WF output, but this framework implicitly accepts the assumption that the uncertainty is homogeneous; thus, this framework does not express differences in confidence in the prediction results depending on the situation.

To characterize these uncertainty representations, two representative PDP performance metrics, the *prediction interval normalized average width* (PINAW) and the *prediction interval coverage probability* (PICP), are introduced here. The PINAW [48] is one of the popular metrics for evaluating the PDP from the viewpoint of the average width of the percentile range γ in the results. In this experiment, we evaluated PINAW for several types of percentile range, γ, for each timeslot; when the PINAW is small, the interval prediction range γ derived from the PDP is tight on average for the timeslot t. The PICP [48] is another popular metric for evaluating the prediction intervals derived from the PDP.



(a) PINAW under $\gamma = 50$. (b) PINAW under $\gamma = 90$. (c) PINAW under $\gamma = 98$.

FIGURE 5. PINAWs under $\gamma \in \{50, 90, 98\}$[%]; "PDP" indicates the framework introduced in Section III-B, and "EPDF-based" indicates the framework used in CWPSH [1].



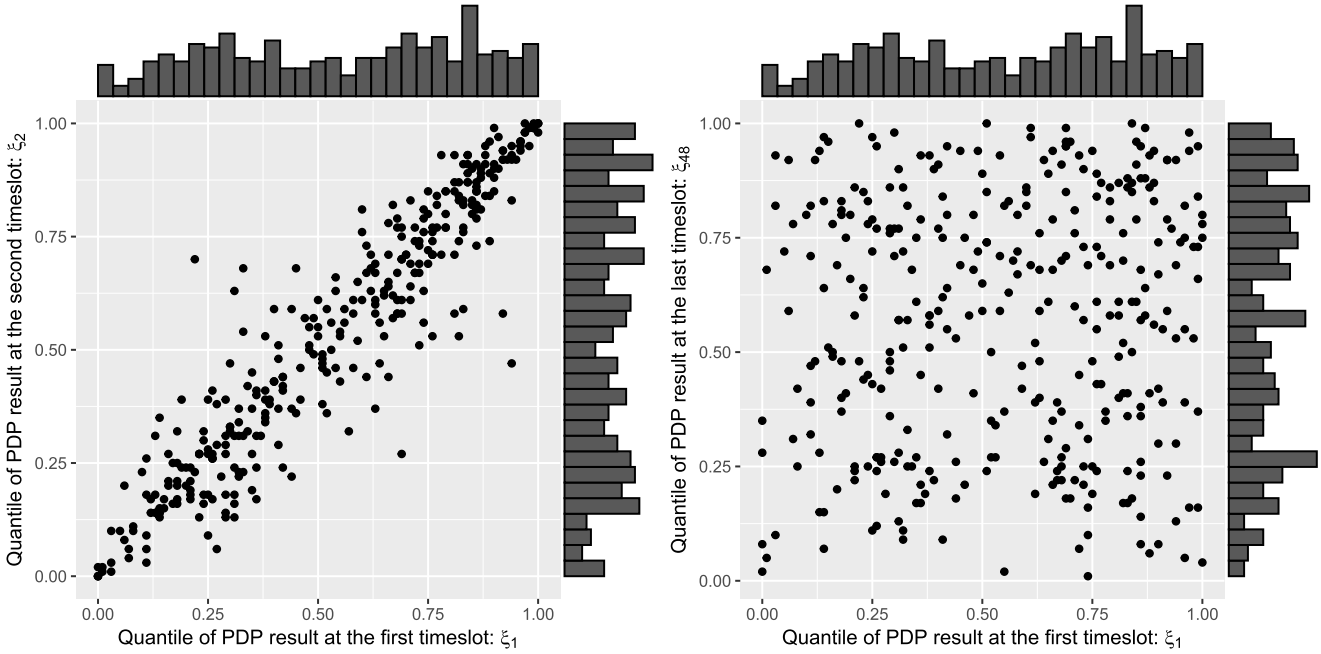
(a) PICP under $\gamma = 50$. (b) PICP under $\gamma = 90$. (c) PICP under $\gamma = 98$.

FIGURE 6. PICPs under $\gamma \in \{50, 90, 98\}$[%]; the dotted lines show the ideal coverage probability for the given γ.

We introduce this metric for the given quantile range γ of the timeslot t to evaluate the mean coverage of the actual WF output; when the PICP is close to γ, the interval predictions derived from the PDP results appropriately cover the realized WF output as expected, so that the uncertainties represented in the PDP model are elaborate.

Figure 5 shows the PINAW of each timeslot evaluated under $\gamma \in \{50, 90, 98\}$[%], focusing on the target timeslots, $\mathcal{T}_0$, for the operation of the results derived for the evaluation period. The homogeneity of the uncertainty representation of the single EPDF-based approach employed in the CWPSH method can be seen in that the PINAWs do not change in any timeslot under each γ. Meanwhile, the PDP introduced in this study shows a different PINAW in each timeslot. The results also suggest that under $\gamma = 50$[%], the two approaches do not differ significantly. Still, under $\gamma = 90$ and 98[%], the PDP framework introduced in this study tends to exhibit a relatively tighter uncertainty representation. Figure 6 shows the results of PICP under $\gamma \in \{50, 90, 98\}$.The PDP framework shows almost ideal average coverage probability for the target timeslots under all γ; however, the single EPDF-based approach shows far from ideal coverage probability, especially for large γ.

On the other hand, Fig. 7 shows the characteristics of the time-dependence of the uncertainties in prediction results. Figure 7(a) shows the quantiles corresponding to the observed WF output obtained on the cumulative probability distribution derived from the respective PDP results focused on the first timeslot of $\mathcal{T}_0$ (i.e., 00:00–00:30 of the next day) and the next timeslot (00:30–01:00 of the next day) in the evaluation period. Note that marginal histograms of ξ_1 and ξ_2 shown in the figure are the *rank histogram* [43], often used to assess the goodness of the PDP results qualitatively. While neither of these histograms is consistent with a perfectly uniform distribution, they do not show a considerable distributional bias, suggesting that the PDP provides a relatively good representation of the uncertainty in the WF output in a way that adequately reflects the actual behavior of WF output. However, the scatter plot shown in the figure suggests that there is a clear dependence between these timeslots; this implies that when the WF output is observed at the tails of the PDP result in one timeslot, the same trend is observed in the next timeslot. On the contrary, Fig. 7(b)



(a) Relationship b/w the first and second timeslots in $\mathcal{T}_0$.

(b) Relationship b/w the first and last timeslots in $\mathcal{T}_0$.

FIGURE 7. Relationships of uncertainties in PDPs between different timeslots. $\xi_t = \Pr(W_t < w_t^*)$ shows the marginal cumulative distribution of the PDP result respective to the observation w_t^*.

shows the relationship between quantiles focused on the first and the last timeslot (23:30–24:00 of the next day) in $\mathcal{T}_0$, suggesting that this dependency is fragile. As illustrated in these examples, the uncertainty in the WF output represented by the PDP can result in various types of dependencies among timeslots.

Figure 8 shows the surface plot of the following bivariate probability density,

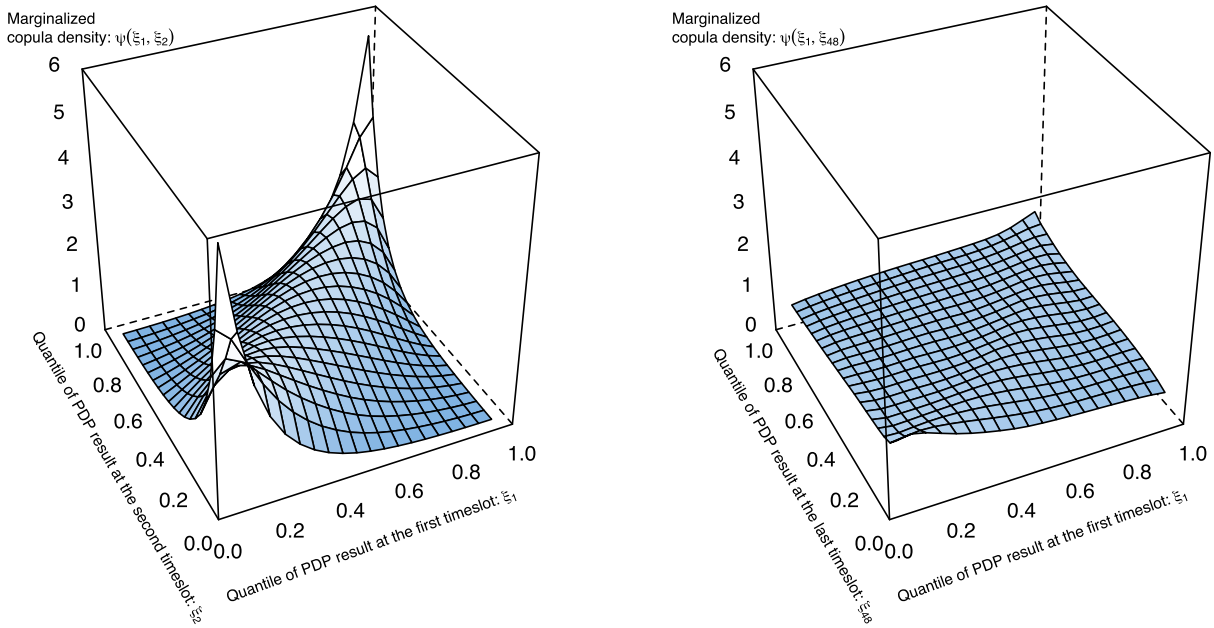
$$\psi(\xi_i, \xi_j) = \int \cdots \iint \cdots \iint \cdots \int \psi(\xi_1, \ldots, \xi_{|\mathcal{T}_+|}) \, d\xi_1 \ldots d\xi_{i-1} d\xi_{i+1} \ldots d\xi_{j-1} d\xi_{j+1} \ldots d\xi_{|\mathcal{T}_+|}, \tag{37}$$

which is obtained by marginalizing the training result of the copula density function, $\psi(.)$. Figure 8(a) suggests that the copula density function has properly learned the strong temporal dependency between the first and the second timeslots in $\mathcal{T}_0$ shown in Fig. 7(a). The training result well represents the inherent property of PDP results that if the observed value deviates significantly from the PDP in the direction of over-/under-estimation in the first timeslot, they tend to deviate in the same direction in the following timeslot. Meanwhile, the training result shown in Fig. 8(b) suggests that no strong dependency was identified between the first and last timeslots of $\mathcal{T}_0$, which is consistent with the characteristics shown in Fig. 7(b). Thus, the approach introduced in Section III-D incorporates these time dependency characteristics in a data-centric manner and provides important information for a detailed BG operational plan with PSHG.

C. EVALUATION OF BG SCHEMES

Now, to provide an overview of the annual performance trends of each method, we present the comparison results focusing on the proposed BG methods, which employ scenarios generated under $J = 10$ and $S = 10$.

Figure 9 shows the results of the daily average revenues. The comparison between the results of the *indep.* and *naive* approaches suggests the effectiveness of the fundamental BG operation configured with the WF and PSHG. Meanwhile, the *CWPSH*, *HWJB*, and the *proposed* BG schemes exhibit even better results than those of the *indep.* and *naive* schemes in terms of realized revenue; the results suggest the effectiveness of the cooperative BG planning scheme of operational schedules for addressing the impact of uncertainties in the WF output. The comparison between *CWPSH* and the *HWJB* suggests that considering the *heterogeneity* in the BG operation scheduling improves profitability, and the comparison between *HWJB* and the *proposed* scheme suggests that considering the *time-dependence* in capacity constraints further improves profitability. Table 3 shows the breakdown of daily average revenues. Although the proposed methods increase the imbalance charges compared with the *indep.* and *naive* approaches, the proposed BG schemes achieve more enormous revenues owing to the increase in power selling revenue and decrease in power purchase expenditure. In particular, modeling temporal dependencies contributes to increasing electricity sales while reducing power purchase expenditure. Figure 10 shows the frequency with which the PSHG water level reached its operating limit of 30-min timeslots during the control phase of compensating for deviations of the WF output from the plan in the BG



(a) Time-dependency training result b/w the first and second timeslots in $\mathcal{T}_0$.

(b) Time-dependency training result b/w the first and last timeslots in $\mathcal{T}_0$.

FIGURE 8. Training results of time-dependency between different timeslots.

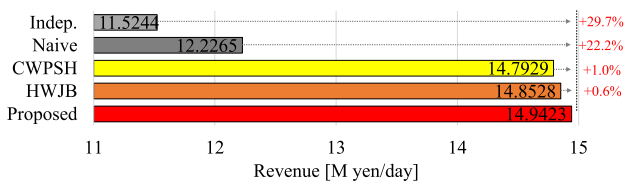


FIGURE 9. Daily average revenue.

TABLE 3. Breakdown of daily average revenue.

Description	*Indep.*	*Naive*	*CWPSH*	*HWJB*	*Proposed*
Electricity sales	19.4758	19.9247	39.9500	36.4585	38.0946
Power purchase	-2.8499	-2.7515	-0.9173	-2.1480	-1.2762
Imbalance charge	-5.1015	-4.9468	-24.2398	-19.4577	-21.8761
Total revenue	11.5244	12.2265	14.7929	14.8528	14.9423

Unit: million yen.

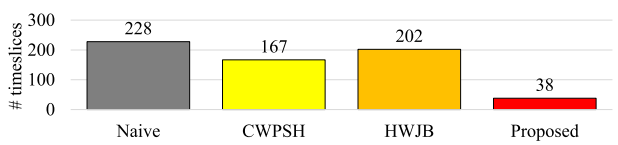


FIGURE 10. Frequency of capacity constraints violation.

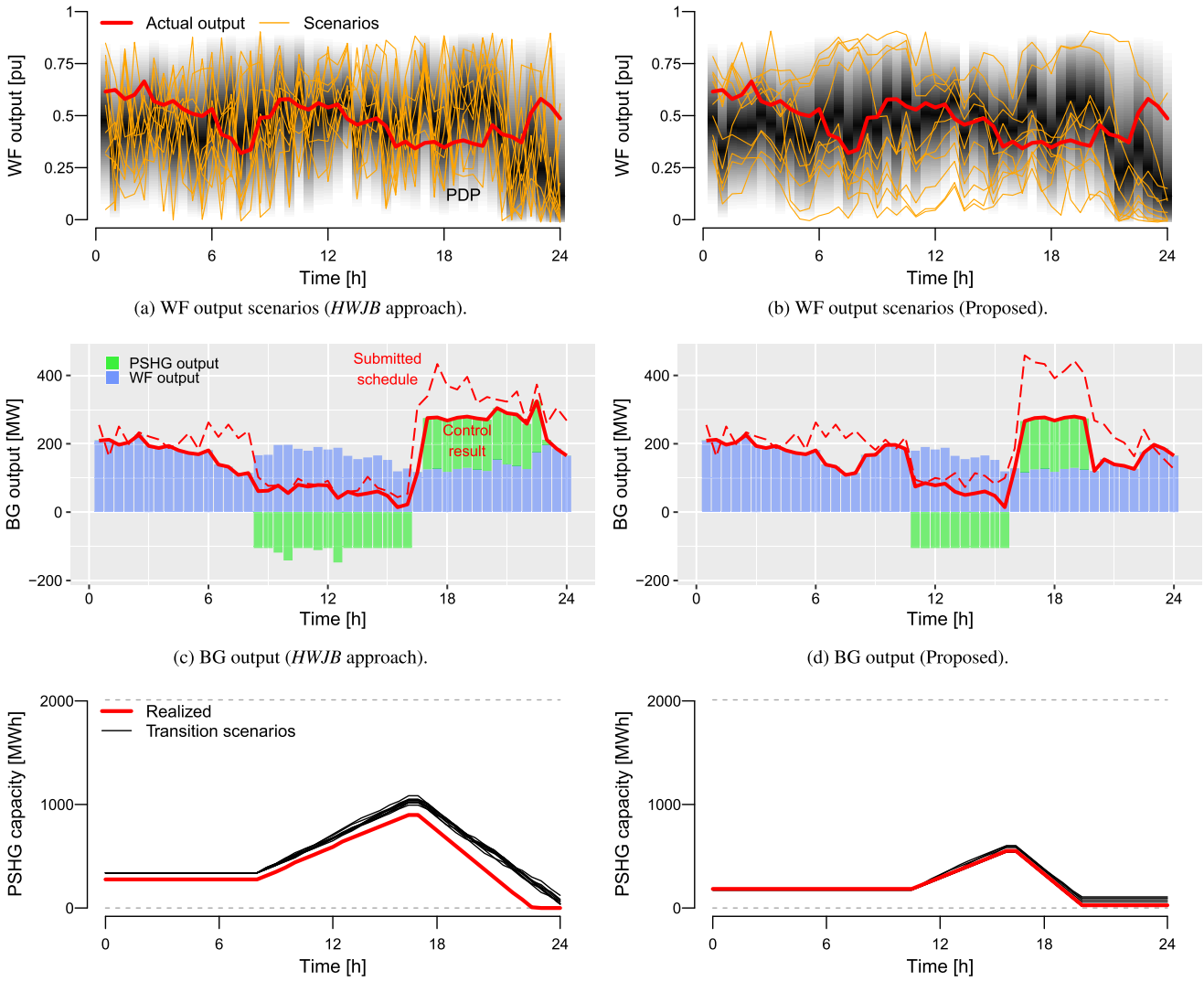
frameworks. The proposed method shows significantly fewer violations than the *HWJP* and *CWPSH* approaches. The results suggest that considering the temporal dependencies in the WF output contributes to the appropriate capacity management of the PSHG for the BG performance; this benefit contributes to realizing the operational schedules reducing the amount of additional power purchased while achieving high profitability.

Figure 11 shows the operation results on a representative day. Figures 11(a) and (b) well characterize the WF output scenarios used in the proposed scheme. Without considering the temporal dependencies, the outputs tend to have unnatural variability as shown in Fig. 11(a); however, explicit consideration of the temporal dependencies encourages the derivation of natural scenarios as shown in Fig. 11(b). Mainly due to the nature of the scenarios used, differences in the BG operational schedules become apparent, as shown in Figs. 11(c) and 11(d). In particular, the plan derived by the HWJB approach shown in Fig. 11(c) tends to anticipate a longer operating time for PSHG compared to that of the proposed approach shown in Fig. 11(d). Meanwhile, Fig. 11(e) shows that the water level transition in the control phase deviates significantly from the transition scenarios if the temporal dependencies are not considered, owing to compensation for the WF output, and reaches the operation limit at the end of the day. However, considering temporal dependencies keeps that transitions in the control phase follow the scenarios derived during scheduling and that the water storage capacity is operated within limits, as shown in Fig. 11(f). These results demonstrate that consideration of the time-dependent structure of uncertainty in WF output adequately captures the impact of potential reservoir transitions on PSHG operational planning.

D. DISCUSSION ON THE PROPOSED OPTIMIZATION SCHEME

1) COMPUTATIONAL BURDEN

The iterative discretization procedure described in Algorithm 2 was introduced to efficiently achieve expected revenue optimization, which is difficult to implement directly. To discuss the effectiveness of the proposed iterative discretization procedure, we evaluated the derived expected revenues and computational times compared to the results



(a) WF output scenarios (*HWJB* approach).

(b) WF output scenarios (Proposed).

(c) BG output (*HWJB* approach).

(d) BG output (Proposed).

(e) Water level (*HWJB* approach).

(f) Water level (Proposed).

FIGURE 11. Operational results of the *HWJB* [2] approach and the proposed approach on a representative day.

in the case where the optimization was performed directly by discretizing at high resolution for all the target timeslots without an iterative procedure (i.e., the initial discretization parameter N^0 in Algorithm 2 was set to $N^0 = \bar{N} = 320$); the comparison was conducted focusing on one typical week. Figure 12 shows the expected revenues and the maximizers of upper/lower bounds given in Eqs. (32) and (34) derived for each target date. The result suggests that the surrogate functions introduced in Section III-E give tight bounds on the objective function of the expected revenue optimization problem. The result also shows that the proposed iterative discretization approach achieves a performance comparable to optimization with high resolution discretization for all the target timeslots.

Meanwhile, Fig. 13 shows the computational times required to achieve the optimization. The results show that the proposed iterative discretization optimization approach effectively reduces, in most cases, the computational burden required to achieve the expected revenue optimization of comparable quality to that based on high resolution discretization. In particular, these results imply that the proposed procedure works effectively in situations like the one introduced in Section III-A, where day-ahead scheduling is required by 10:00 based on new information accessible at 8:00 in the real world. Thus, in terms of computational requirements, it suggests that the proposed iterative optimization scheme with surrogate functions provides a sufficiently practical and efficient solution.

2) SENSITIVITY TO SCENARIO SELECTION

Even if the proposed BG scheme successfully accounts for temporal dependencies, the ratio of the number of adopted scenarios to that of generated scenarios, which controls the relative plausibility, may affect the operational results. In particular, the sensitivity of the optimality to the number of scenarios J, which can have an impact on the

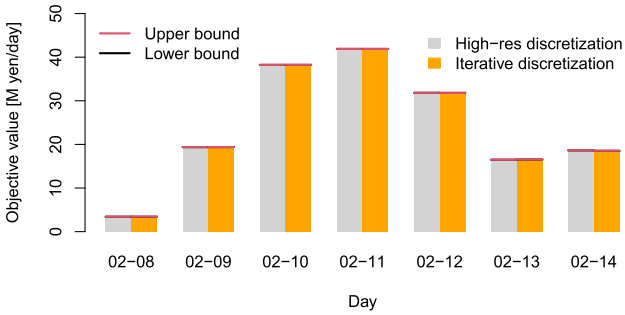


FIGURE 12. Derived upper/lower bounds for representative days.

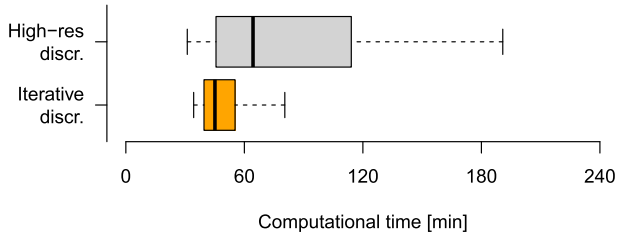


FIGURE 13. Computational time required for planning.

BG performance without having a significant impact on the computational burden of the optimization, provides an important perspective when tuning the proposed method. Therefore, we evaluated the impact of the scenario selection guideline related to the PSHG capacity constraints on the BG performance. We focus on the following four cases with different numbers of scenarios: $J \in \{10, 20, 100, 1000\}$ under $S = 10$; these cases represent the conditions for PSHG capacity chance constraints, i.e., 100%, 50%, 10%, and 1%. The other experimental conditions are the same as those described in Section IV-A.

Figure 14 shows the annual simulation results of daily average revenues, which indicate that the respective revenues are comparable. The results show that the proposed method does not necessarily require the number of scenarios to be carefully tuned in selecting the WF output scenarios to calculate capacity constraints, and the method tends to be robust and stable in determining a reasonable schedule. In determining the scenario selection ratio to be used in a real-world situation, one possible approach is to dynamically tune the ratio of S to J with the highest revenue based on simulation verification using historical data, e.g., for the most recent week. However, our results indicate that revenue achievement is not highly sensitive to scenario selection ratios. Thus, the experimental results imply that an appropriate operational performance can be expected even if the ratio is given in an ad-hoc manner.

V. CONCLUSION

This study focused on a BG scheme consisting of an WF and PSHG and proposed an operational scheduling framework for maximizing expected revenues. The PDP scheme for WF output and copula-based scenario generation scheme were introduced to reflect the impact of uncertainties in WF output on the PSHG water level transition; scenarios generated in our scheme particularly consider the heterogeneity and temporal dependencies in the PDP results of WF outputs. The proposed scheme efficiently derives the formulated optimization problem by iteratively updating surrogate functions while tightening the upper/lower bound evaluations. The results of the annual numerical experiments show the following findings:

- The proposed method achieves more significant revenues than comparative methods by appropriately managing the water level of PSHG.
- In particular, the effects of accounting for the heterogeneity and temporal dependency in the uncertainty of WF output were quantitatively evaluated and experimentally demonstrated to have a 29.8 [%] revenue improvement potential compared to the independent operation of WF and PSHG and a 22.2 [%] improvement to the naive BG operation.

We also believe this effort will provide insight into effectively applying the state-of-the-art PDP scheme for renewable energy sources to real-world energy management. Moreover, the proposed iterative optimization framework is not restricted by the class of probability distributions assumed in the PDP of WF output; therefore, it is worth mentioning that even PDP results assuming nonparametric distributions can be directly utilized.

Our numerical experiments show that the proposed method is effective even when a naive prediction framework is used for the electricity price and imbalance charge; however, the impact of the prediction scheme's characteristics on the profitability of the operational results is an exciting topic. Another important topic, which was not detailed in this paper, is to further improve the practicality of operation in a situation where WFs and PSHG are spatially distant and form a BG scheme, considering the possibility that the operational feasibility may be affected by the operational status of the power system (e.g., congestion in terms of transmission line capacity) and the compatibility with the grid frequency control capability that PSHGs are currently responsible for. In addition to the day-ahead market addressed in this study, using other types of markets is crucial in the BG operation of renewable energy. We plan to continue our research with these open issues.

APPENDIX A
REGULAR VINE COPULA

The copula introduced in Section III-D enables us to represent the joint probability of multiple random variables using the marginal probabilities (i.e., the PDP results of the WF output for each timeslot) and copula density function $\psi(\xi_1, \ldots, \xi_{|\mathcal{T}_+|})$; the copula has been used in the context of modeling the dependence of power generation between multiple WFs in related fields [47], [49], [50] and is known to be a useful idea for representing the temporal dependence of the WF output behavior among timeslots [33]. In particular, the R-vine copula adopted in this study expresses this multidimensional copula density function using the nested

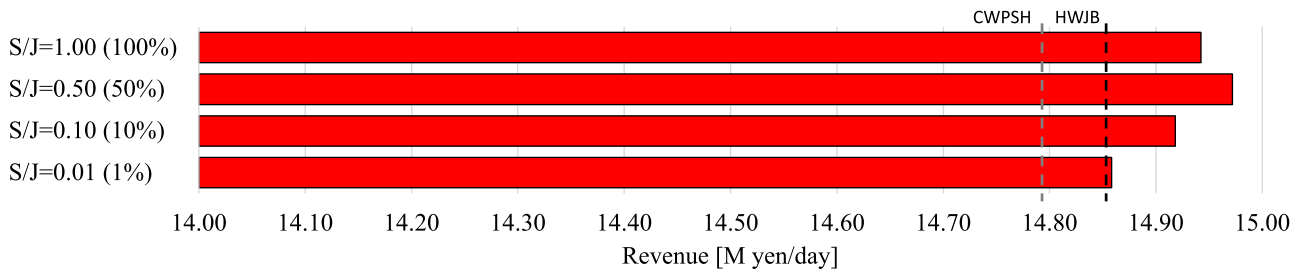


FIGURE 14. Daily average revenue under various conditions for chance constraints in the proposed method.

combination of parametric bivariate copula functions (i.e., pair-copula construction [35]), which are rich in expressive power and have been discussed extensively.

Let $\{\mathcal{G}_i; i \in \{1, \ldots, |\mathcal{T}_+| - 1\}\}$ be a set of undirected tree graphs defining the structure of the R-vine copula; $\mathcal{G}_1$ has nodes $\mathcal{T}_+$ and edges $\mathcal{E}_1 = \{(a, b); a, b \in \mathcal{T}_+\}$, and $\mathcal{G}_i (i = 2, \ldots, |\mathcal{T}_+| - 1)$ has nodes $\mathcal{E}_{i-1}$ and edges $\mathcal{E}_i = \{(\mathcal{A}, \mathcal{B}); \mathcal{A}, \mathcal{B} \in \mathcal{E}_{i-1}\}$. By introducing the idea of the R-vine copula, we represent the copula density function, $\psi(\xi_1, \ldots, \xi_{|\mathcal{T}_+|})$, shown in Eq. (23) as follows[9]:

$$\psi(\xi_1, \ldots, \xi_{|\mathcal{T}_+|}) = \prod_{j=1}^{|\mathcal{T}_+|-1} \prod_{(\mathcal{A},\mathcal{B})\in\mathcal{E}_j} \psi_{a,b|\mathcal{A}\cap\mathcal{B}}(\xi_{a|\mathcal{A}\cap\mathcal{B}}, \xi_{b|\mathcal{A}\cap\mathcal{B}}), \tag{38}$$

where a and b indicate the conditioned variable indices $a = \mathcal{A} \setminus (\mathcal{A} \cup \mathcal{B})$ and $b = \mathcal{B} \setminus (\mathcal{A} \cup \mathcal{B})$, respectively, $\psi_{a,b|\mathcal{A}\cap\mathcal{B}}$ indicates the bivariate copula function representing the relationship between w_a and w_b under the conditioning set $\mathcal{A}\cap\mathcal{B}(\subset \mathcal{T}_+)$, and $\xi_{a|\mathcal{A}\cap\mathcal{B}}$ and $\xi_{b|\mathcal{A}\cap\mathcal{B}}$ indicate the conditional cumulative distributions, i.e.,

$$\xi_{a|\mathcal{A}\cap\mathcal{B}} = \Pr\left(W_a < w_a \,\middle|\, \bigcup_{t\in\mathcal{A}\cap\mathcal{B}} w_t\right),$$
$$\xi_{b|\mathcal{A}\cap\mathcal{B}} = \Pr\left(W_b < w_b \,\middle|\, \bigcup_{t\in\mathcal{A}\cap\mathcal{B}} w_t\right). \tag{39}$$

In this study, the popular classes of parametric copulas [46] (i.e., Independent, Gaussian, Student t, Clayton, Gumbel, and Joe copulas) were introduced as candidates of the bivariate copula density functions, $\psi_{a,b|\mathcal{A}\cap\mathcal{B}}$, and these models were fitted based on the maximum likelihood estimation. Note that the representation of the joint probability as an R-vine depends on the given graph set $\{\mathcal{G}_i\}$; therefore, multiple structural representations generally exist. In the numerical experiments conducted in this study, we employed the forward tree selection scheme [51], and the best combination of parametric bivariate copulas was selected according to the Akaike's information criterion [52].

[9]For a more detailed explanation, please see [35].

APPENDIX B
DERIVATION OF SURROGATE FUNCTIONS

The surrogate functions introduced in Section III-E are derived using the discrete approximations of the expectation operation in the original objective function, Eqs. (13), under the PDP results of wind power generation. We focus on the concavity of $\hat{r}_t(w_t \mid \theta_t)$ shown in Eq. (14) for a fixed θ_t. Using Jensen's inequality, the original objective function, Eq. (13), has the following property under the given partition set $\{\mathcal{N}_t\}$,

$$\begin{aligned}
F(\{\theta_t\}) &= \sum_{t\in\mathcal{T}_+} \mathbb{E}\left[\hat{r}_t(W_t \mid \theta_t)\right] \\
&= \sum_{t\in\mathcal{T}_+}\sum_{n\in\mathcal{N}_t} \underbrace{\int_{\underline{w}_t^n}^{\bar{w}_t^n} \Pr(w_t)dw_t}_{p_t^n} \mathbb{E}\left[\hat{r}_t(W_t \mid \theta_t) \mid \mathcal{W}_t^n\right] \\
&\le \sum_{t\in\mathcal{T}_+}\sum_{n\in\mathcal{N}_t} p_t^n \hat{r}_t\left(\int_{\underline{w}_t^n}^{\bar{w}_t^n} \frac{w_t}{p_t^n} \Pr(w_t)dw_t \mid \theta_t\right) \\
&= \sum_{t\in\mathcal{T}_+}\sum_{n\in\mathcal{N}_t} p_t^n \hat{r}_t\left(\frac{\tilde{w}_t^n}{p_t^n} \mid \theta_t\right) \\
&= \bar{F}(\{\theta_t\}; \{\mathcal{N}_t\}),
\end{aligned} \tag{40}$$

where $\mathbb{E}[.|\mathcal{W}_t^n]$ indicates the conditional expectation under a given subregion from $\underline{w}_t^n$ to $\bar{w}_t^n$. Note that the partitioning policy shown in Eqs. (30) and (31) ensures that $p_t^n = \frac{1}{N_t}$. Similarly, another surrogate function representing the lower bound of the original objective function is derived using the Edmundson-Madansky bound:

$$\begin{aligned}
F(\{\theta_t\}) &= \sum_{t\in\mathcal{T}_+} \mathbb{E}\left[\hat{r}_t(W_t \mid \theta_t)\right] \\
&= \sum_{t\in\mathcal{T}_+}\sum_{n\in\mathcal{N}_t} \int_{\underline{w}_t^n}^{\bar{w}_t^n} \hat{r}_t(w_t \mid \theta_t) \Pr(w_t)dw_t \\
&\ge \sum_{t\in\mathcal{T}_+}\sum_{n\in\mathcal{N}_t} \left\{\frac{\bar{w}_t^n - \tilde{w}_t^n}{\bar{w}_t^n - \underline{w}_t^n}\hat{r}(\underline{w}_t^n|\theta_t) + \frac{\tilde{w}_t^n - \underline{w}_t^n}{\bar{w}_t^n - \underline{w}_t^n}\hat{r}(\bar{w}_t^n|\theta_t)\right\} \\
&= \underline{F}(\{\theta_t\}; \{\mathcal{N}_t\}).
\end{aligned} \tag{41}$$

REFERENCES

[1] Y. Li, F. Wu, J. Li, Y. Yin, Z. Li, and L. Ai, "Chance-constrained energy management for pumped storage hydropower plant to compensate for wind power uncertainties," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Jul. 2021, pp. 1–5.

[2] A. Cerejo, S. J. P. S. Mariano, P. M. S. Carvalho, and M. R. A. Calado, "Hydro-wind optimal operation for joint bidding in day-ahead market: Storage efficiency and impact of wind forecasting uncertainty," *J. Mod. Power Syst. Clean Energy*, vol. 8, no. 1, pp. 142–149, Jan. 2020.

[3] Y. Fujimoto, Y. Takahashi, and Y. Hayashi, "Alerting to rare large-scale ramp events in wind power generation," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 55–65, Jan. 2019.

[4] M. Gržanić and T. Capuder, "Coordinated scheduling of renewable energy balancing group," *Int. J. Electr. Power Energy Syst.*, vol. 125, Feb. 2021, Art. no. 106555.

[5] L. Koltermann, K. Jacqué, J. Figgener, S. Zurmühlen, and D. U. Sauer, "Balancing group deviation & balancing energy costs due to the provision of frequency containment reserve with a battery storage system in Germany," *Int. J. Electr. Power Energy Syst.*, vol. 142, Nov. 2022, Art. no. 108327.

[6] M. Inagaki, A. Kaneko, Y. Fujimoto, Y. Hayashi, and S. Minotsu, "Balancing group operation for wind power plants and a pumped-storage hydro generator to maximize expected revenue in consideration of wind power output uncertaintiy," in *Proc. 21st Wind Sol. Integr. Workshop (WIW)*, 2022, pp. 650–655.

[7] Y. Fujimoto, H. Ishii, and Y. Hayashi, "Designing sustainable smart cities: Cooperative energy management systems and applications," *IEEJ Trans. Electr. Electron. Eng.*, vol. 15, no. 9, pp. 1256–1270, Sep. 2020.

[8] Z. Zhang, Y. Zhang, Q. Huang, and W.-J. Lee, "Market-oriented optimal dispatching strategy for a wind farm with a multiple stage hybrid energy storage system," *CSEE J. Power Energy Syst.*, vol. 4, no. 4, pp. 417–424, Dec. 2018.

[9] S. Mohammadi, S. Soleymani, and B. Mozafari, "Scenario-based stochastic operation management of MicroGrid including wind, photovoltaic, micro-turbine, fuel cell and energy storage devices," *Int. J. Electr. Power Energy Syst.*, vol. 54, pp. 525–535, Jan. 2014.

[10] M. Farrokhseresht, N. G. Paterakis, M. Gibescu, and J. G. Slootweg, "Participation of a combined wind and storage unit in the day-ahead and local balancing markets," in *Proc. 15th Int. Conf. Eur. Energy Market (EEM)*, Jun. 2018, pp. 1–5.

[11] K. Das, A. L. T. P. Grapperon, P. E. Sørensen, and A. D. Hansen, "Optimal battery operation for revenue maximization of wind-storage hybrid power plant," *Electr. Power Syst. Res.*, vol. 189, Dec. 2020, Art. no. 106631.

[12] A. A. Thatte, L. Xie, D. E. Viassolo, and S. Singh, "Risk measure based robust bidding strategy for arbitrage using a wind farm and energy storage," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2191–2199, Dec. 2013.

[13] R. Nourollahi, B. Mohammadi-Ivatloo, and K. Zare, "Hybrid stochastic/robust offering strategy for coordinated wind power and compressed air energy storage in multielectricity markets," *IEEE Syst. J.*, vol. 16, no. 1, pp. 977–984, Mar. 2022.

[14] M. Ito, Y. Fujimoto, M. Mitsuoka, H. Ishii, and Y. Hayashi, "Control methods for an energy storage system when wind power output deviates from grid code," *J. Int. Council Electr. Eng.*, vol. 7, no. 1, pp. 159–165, Jan. 2017.

[15] A. A. Salimi, A. Karimi, and Y. Noorizadeh, "Simultaneous operation of wind and pumped storage hydropower plants in a linearized security-constrained unit commitment model for high wind energy penetration," *J. Energy Storage*, vol. 22, pp. 318–330, Apr. 2019.

[16] M. E. Nazari and M. M. Ardehali, "Optimal coordination of renewable wind and pumped storage with thermal power generation for maximizing economic profit with considerations for environmental emission based on newly developed heuristic optimization algorithm," *J. Renew. Sustain. Energy*, vol. 8, no. 6, Nov. 2016, Art. no. 065905.

[17] J. Garcia-Gonzalez, R. M. R. de la Muela, L. M. Santos, and A. M. Gonzalez, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 460–468, May 2008.

[18] A. K. Varkani, A. Daraeepour, and H. Monsef, "A new self-scheduling strategy for integrated operation of wind and pumped-storage power plants in power markets," *Appl. Energy*, vol. 88, no. 12, pp. 5002–5012, Dec. 2011.

[19] M. Parastegari, R.-A. Hooshmand, A. Khodabakhshian, and Z. Forghani, "Joint operation of wind farms and pump-storage units in the electricity markets: Modeling, simulation and evaluation," *Simul. Model. Pract. Theory*, vol. 37, pp. 56–69, Sep. 2013.

[20] A. A. S. de la Nieta, J. Contreras, and J. P. S. Catalão, "Optimal single wind hydro-pump storage bidding in day-ahead markets including bilateral contracts," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1284–1294, Jul. 2016.

[21] M. E. Nazari and M. M. Ardehali, "Optimal bidding strategy for a GENCO in day-ahead energy and spinning reserve markets with considerations for coordinated wind-pumped storage-thermal system and CO_2 emission," *Energy Strategy Rev.*, vol. 26, Nov. 2019, Art. no. 100405.

[22] M. He, B. Cao, W. He, X. Li, and X. Wu, "Rolling optimization of wind power and pumped hydro storage system based on W-IGDT," in *Proc. IEEE 5th Int. Electr. Energy Conf. (CIEEC)*, May 2022, pp. 4314–4319.

[23] M. Majidi, B. Mohammadi-Ivatloo, and A. Soroudi, "Application of information gap decision theory in practical energy problems: A comprehensive review," *Appl. Energy*, vol. 249, pp. 157–165, Sep. 2019.

[24] X. Jin, Z. Xu, and W. Qiao, "Condition monitoring of wind turbine generators using SCADA data analysis," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 202–210, Jan. 2021.

[25] J. G. Powers et al., "The weather research and forecasting (WRF) model: Overview, system efforts, and future directions," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 8, pp. 1717–1737, 2017.

[26] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.

[27] Y.-K. Wu, P.-E. Su, T.-Y. Wu, J.-S. Hong, and M. Y. Hassan, "Probabilistic wind-power forecasting using weather ensemble models," *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5609–5620, Nov. 2018.

[28] Y.-K. Wu, Y.-C. Wu, J.-S. Hong, L. H. Phan, and Q. D. Phan, "Probabilistic forecast of wind power generation with data processing and numerical weather predictions," *IEEE Trans. Ind. Appl.*, vol. 57, no. 1, pp. 36–45, Jan. 2021.

[29] Z. Men, E. Yee, F.-S. Lien, D. Wen, and Y. Chen, "Short-term wind speed and power forecasting using an ensemble of mixture density neural networks," *Renew. Energy*, vol. 87, pp. 203–211, Mar. 2016.

[30] D. Nohara, M. Ohba, T. Watanabe, and S. Kadokura, "Probabilistic wind power prediction based on ensemble weather forecasting," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 12151–12156, 2020.

[31] Y. Fujimoto, D. Nohara, Y. Kanno, M. Ohba, T. Kato, and Y. Hayashi, "Probability density prediction of wind farm power generation: Benchmarking natural gradient boosting approach using ensemble weather forecast," *SSRN Electron. J.*, vol. 2022, pp. 1–36, Apr. 2022, doi: 10.2139/ssrn.4097480.

[32] M. Inagaki, A. Kaneko, Y. Fujimoto, Y. Hayashi, and S. Minotsu, "Assessing the operational potential of pumped-storage hydro generators for supporting the grid integration of wind farms," in *Proc. IEEE PES Innov. Smart Grid Technol.-Asia*, Dec. 2021, pp. 1–5.

[33] R. Becker, "Generation of time-coupled wind power infeed scenarios using pair-copula construction," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1298–1306, Jul. 2018.

[34] X. Dui, M. Ito, Y. Fujimoto, Y. Hayashi, G. Zhu, and L. Yao, "Time series model of wind power forecasting error by using beta distribution for optimal sizing of battery storage," *IEEJ Trans. Power Energy*, vol. 139, no. 3, pp. 212–224, 2019.

[35] T. Bedford and R. M. Cooke, "Vines—A new graphical model for dependent random variables," *Ann. Statist.*, vol. 30, no. 4, pp. 1031–1068, Aug. 2002.

[36] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

[37] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. Cham, Switzerland: Springer, 1998.

[38] K. B. Kim, "Receding horizon tracking control for constrained linear continuous time-varying systems," *IEE Proc.-Control Theory Appl.*, vol. 150, no. 5, pp. 534–538, Sep. 2003.

[39] A. Yoshida, J. Yoshikawa, Y. Fujimoto, Y. Amano, and Y. Hayashi, "Stochastic receding horizon control minimizing mean-variance with demand forecasting for home EMSs," *Energy Buildings*, vol. 158, pp. 1632–1639, Jan. 2018.

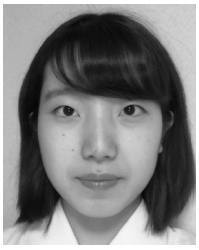
[40] Y. Fujimoto, H. Kikusato, S. Yoshizawa, S. Kawano, A. Yoshida, S. Wakao, N. Murata, Y. Amano, S. Tanabe, and Y. Hayashi, "Distributed energy management for comprehensive utilization of residential photovoltaic outputs," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1216–1227, Mar. 2018.

[41] T. Duan, A. Avati, D. Y. Ding, S. Basu, A. Y. Ng, and A. Schuler, "NGBoost: Natural gradient boosting for probabilistic prediction," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 2690–2700.
[42] A. Hashimoto and H. Hirakuchi, "Enhancement and accuracy evaluation of numerical weather forecasting and analysis system (NuWFAS) for the Hokkaido Island," Central Res. Inst. Electr. Power Ind., Tokyo, Japan, Tech. Rep. N09024, 2010.
[43] G. Candille and O. Talagrand, "Evaluation of probabilistic prediction systems for a scalar variable," *Quart. J. Roy. Meteorolog. Soc.*, vol. 131, no. 609, pp. 2131–2150, Jul. 2005.
[44] G. Sideratos and N. D. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1788–1796, Nov. 2012.
[45] S. Amari, "Natural gradient works efficiently in learning," *Neural Comput.*, vol. 10, no. 2, pp. 251–276, Feb. 1998.
[46] R. B. Nelsen, *An Introduction to Copulas* (Springer Series in Statistics), 2nd ed. Cham, Switzerland: Springer, 2006.
[47] Z. Wang, W. Wang, C. Liu, and B. Wang, "Forecasted scenarios of regional wind farms based on regular vine copulas," *J. Modern Power Syst. Clean Energy*, vol. 8, no. 1, pp. 77–85, Jan. 2020.
[48] I. K. Bazionis and P. S. Georgilakis, "Review of deterministic and probabilistic wind power forecasting: Models, methods, and future research," *Electricity*, vol. 2, no. 1, pp. 13–47, Jan. 2021.
[49] X. Yundai and Y. Yue, "Analysis of aggregated wind power dependence based on optimal vine copula," in *Proc. IEEE PES Innov. Smart Grid Technol. Asia*, Mar. 2019, pp. 1788–1792.
[50] J. Xu, W. Wu, K. Wang, and G. Li, "C-vine pair copula based wind power correlation modelling in probabilistic small signal stability analysis," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 4, pp. 1154–1160, Jul. 2020.
[51] K. Aas, C. Czado, A. Frigessi, and H. Bakken, "Pair-copula constructions of multiple dependence," *Insurance, Math. Econ.*, vol. 44, no. 2, pp. 182–198, Apr. 2009.
[52] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proc. 2nd Int. Symp. Inf. Theory*, 1973, pp. 267–281.

YU FUJIMOTO (Member, IEEE) received the Ph.D. degree in engineering from Waseda University, Tokyo, Japan, in 2007.

He is currently a Professor with the Advanced Collaborative Research Organization for Smart Society (ACROSS), Waseda University. His research interests include machine learning, statistical data analysis, data mining in energy domains, especially for controlling power in smart grids, and statistical prediction of the power fluctuation under the extensive introduction of renewable energy sources.

Prof. Fujimoto is a member of the Information Processing Society of Japan.



MAIKO INAGAKI (Member, IEEE) received the B.Eng. and M.Eng. degrees from Waseda University, Japan, in 2021 and 2023, respectively.

Since 2023, she has been working in the field of electrical equipment design. Her research interest includes the cooperative operation of wind turbines with existing generators (mainly pumped-storage hydro generators) for stable utilization of wind turbines.

AKIHISA KANEKO (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in engineering from Waseda University, Tokyo, Japan, in 2015, 2017, and 2020, respectively.

He is currently an Assistant Professor with Waseda University. His current research interests include the operation and control of distribution and transmission network with renewable energy sources.

Dr. Kaneko is a member of the Institute of Electrical Engineers of Japan.

SHINICHIRO MINOTSU received the B.E. degree in electronic engineering and the M.S. degree in information science and technology from The University of Tokyo, in 2006 and 2008, respectively. He is currently a Power System Engineer with Electric Power Development Company Ltd. (J-POWER). His research interests include the operation, economics, and modeling of power and energy systems. He is a member of the Institute of Electrical Engineers in Japan (IEEJ).

YASUHIRO HAYASHI (Member, IEEE) received the B.Eng., M.Eng., and D.Eng. degrees from Waseda University, Tokyo, Japan, in 1989, 1991, and 1994, respectively.

In 1994, he became a Research Associate with Ibaraki University, Mito, Japan. In 2000, he became an Associate Professor with the Department of Electrical and Electronics Engineering, Fukui University, Fukui, Japan. Since 2009, he has been a Professor with the Department of Electrical Engineering and Bioscience, Waseda University. Since 2010, he has been the Director of the Research Institute of Advanced Network Technology. Since 2014, he has been the Dean of the Advanced Collaborative Research Organization for Smart Society (ACROSS), Waseda University. Since 2022, he has been the Chairperson of ACROSS and the Director of the Waseda Center for a Carbon Neutral Society. His current research interests include the optimization of distribution system operation and forecasting, operation, planning, and control concerned with renewable energy sources and demand response.

Prof. Hayashi is a Board Member of the Institute of Electrical Engineers of Japan.