Data Purification for Improved Power Dispatch Against Renewable Uncertainty

Yanzhi Wang, *Graduate Student Member, IEEE*, Jianxiao Wang, *Senior Member, IEEE*, and Jie Song, *Senior Member, IEEE*

***Abstract*—Advancements in information technology and the exponential growth of data in energy systems present significant potential for intelligent and secure grid operations. However, variability in data quality remains a critical constraint. To address the lack of focus on the role of high-quality data in improving decision-making, this paper proposes a two-stage data purification framework. The first stage employs reinforcement learning-based valuation with refined policy strategies to quantify data quality, providing ranked references for decision-focused filtering in the second stage. Applied to stochastic optimization in wind-integrated unit commitment, the proposed method demonstrates its effectiveness on IEEE 30 and IEEE 118-bus systems by accurately identifying high-quality data and achieving economic benefits under varying uncertainty levels. This work aims to introduce a conceptual framework for data-centric decision-making improvement and provide methodological guidance for both academia and industry.**

***Index Terms*—Data quality, data valuation, data-centric AI, outlier detection, reinforcement learning (RL), Shannon entropy (EN), stochastic optimization, unit commitment (UC), wind power forecasting.**

NOMENCLATURE

Sets and Indices

G	Set of thermal generation units.
T	Time horizon (e.g., 24 h).
U	Set of all units.
W	Set of wind farms.
ϕ_i	Set of transmission lines linking node i.

Variables

ξ	Random vector describe uncertainty.
$P_{i,t}^{AW}(\xi)$	Actual output of wind power in i at time t.
$P_{i,t}^{G}(\xi)$	Power of generator i at time t.
$P_{ij,t}(\xi)$	Power flow of line linked i, j at time t.
$\alpha_{i,t}$	Start-up indicator of generator i at time t.
$\beta_{i,t}$	Shut-down indicator of generator i at time t.
$\upsilon_{i,t}$	Commitment status of generator i at time t.

Parameters and Constants

C^P	Penalty cost coefficient.
C_i^D	Stop cost coefficient of generator i.
C_i^U	Start cost coefficient of generator i.
$F_{c,i}$	Fuel cost function for generator i.
$P_{i,t}^{FW}$	Forecast output of wind power in i at time t.
$P_{i,t}^{W}(\xi)$	Real time output of wind power in i at time t.
$P_{i,t}^{W+}(\xi)$	Positive deviation of power in i in time t.
$P_{i,t}^{W-}(\xi)$	Negative deviation of power in i in time t.
η	Load fluctuation factor.
$P_{i,t}^{L}$	Load in i in time t.
$P_{i,\min}^{G}$	Minimum capacity limit of generator i.
$P_{i,\max}^{G}$	Maximum capacity limit of generator i.
$P_{i,\text{up}}^{G}$	Maximum start-up power of generator i.
$P_{i,\text{down}}^{G}$	Maximum shut-down power of generator i.
$P_{ij,\min}$	Minimum flow limit of line linked i, j.
$P_{ij,\max}$	Maximum flow limit of line linked i, j.
R_i^{down}	Ramp-down rate limit of generator i.
R_i^{up}	Ramp-up rate limit of generator i.
$S_{i,t}(\xi)$	Shortage of wind power output in i.
T_i^{on}	Minimum up time of generator i.
T_i^{off}	Minimum off time of generator i.

I. INTRODUCTION

WITH the advent of smart grids and advancements in information technology, the energy sector has experienced an exponential surge in data generation [1]. Big data technologies capture and summarize complex system patterns, thereby enhancing the security and efficiency of grid operations [2]. For instance, learning-based models address the uncertainty associated with renewable energy by providing more accurate output predictions, enabling more efficient power dispatch [3].

However, most studies focus on the efficient and intelligent development of data-driven models while overlooking the impact of front-end data quality variations on subsequent decision-making. In fact, not all data contribute equally to modeling: Critical datasets that uncover underlying patterns and support

Received 15 June 2024; revised 6 December 2024 and 18 January 2025; accepted 12 February 2025. This work was supported by the National Natural Science Foundation of China under Grant 724B2002 and Grant 72422015 and in part by the Peking University - BHP Carbon and Climate Wei-Ming Ph.D. Scholars Program under Grant WM202407. Paper no. TII-24-2971. *(Corresponding author: Jie Song.)*

Yanzhi Wang and Jie Song are with the Department of Industrial Engineering and Management, College of Engineering, Peking University, Beijing 100871, China (e-mail: yanzhiwang@pku.edu.cn; jie.song@pku.edu.cn).

Jianxiao Wang is with the National Engineering Laboratory for Big Data Analysis and Applications, Peking University, Beijing 100871, China (e-mail: wangjx@pku.edu.cn).

Digital Object Identifier 10.1109/TII.2025.3545092

accurate parameter weighting are invaluable, whereas noisy and biased data can undermine model precision [4]. This highlights the importance of data purification, a data-centric conceptual approach that incorporates data valuation and filtering. By identifying and extracting high-quality data, it significantly enhances both model accuracy and decision-making reliability [5].

Current data purification methods can be broadly categorized into three primary types. The first type employs transaction-based and sharing mechanisms to assign value to datasets, facilitating the selection of high-value data. For instance, Agarwal et al. [6] proposed a mathematical model for a two-sided market mechanism to price training data. This model was applied in renewable energy source (RES) forecasting tasks, where participants who chose higher value datasets achieved greater profits [7]. Similarly, other researchers have proposed data valuation frameworks tailored to specific applications. Examples include a two-stage electricity market model aimed at reducing uncertainty [8] and cooperative game theory-based methods for evaluating dataset contributions in ensemble load models [9]. These advancements have improved data market practices by refining valuation and filtering methods to better identify high-quality datasets. However, their reliance on market mechanisms limits their applicability to broader system optimization and direct decision-making improvement.

The second category emphasizes general statistical metrics for illustrating data valuation. As an unsupervised data preprocessing method, outlier detection is widely recognized as an effective approach to enhance time-series forecasting performance [10]. From a statistical perspective, advanced metrics, such as Shannon entropy (EN) have been employed to quantify information value, with higher EN datasets, such as those related to photovoltaic systems, demonstrating improved accuracy in power system planning and operations [11]. Similarly, Wasserstein distance has been utilized to evaluate the value of multisource datasets in distributionally robust optimization [12]. However, these statistical valuation methods cannot accommodate task-specific supervised adjustments, limiting their general applicability to unbiased data valuation across diverse scenarios.

To address the limitations of statistical valuation methods, the third category of data purification approaches employs supervised algorithms to enhance the trustworthiness of data-driven models [13]. By conceptualizing data as elements in a cooperative game, data Shapley quantifies the average marginal contribution of each data point by iterating over all subsets that exclude it [14]. To improve computational efficiency and generalizability, researchers have introduced KNN-based approximations [15] and distributional extensions [16]. Building on these advancements, more flexible learning-based valuation methods have been developed to accommodate heterogeneous data analytics. Leveraging deep learning, these methods provide unbiased value estimates that are transferable across domains [17]. Building on this, wind power data integrated with multisource meteorological features has been successfully purified, achieving a validated accuracy improvement of 7.69% [18]. However, these methods remain predominantly focused on prediction tasks and do not address the inherent inconsistency between norm-based loss metrics (e.g., MSE, MAPE) and decision-making costs (e.g., optimization objectives). In two-stage power grid scheduling, for example, under-forecasting day-ahead loads can result in unnecessary unit start-ups and surplus power purchases, while over-forecasting leads to costly balancing power and suboptimal dispatch decisions [19]. This mismatch highlights a key limitation of current data purification approaches, as minimizing prediction loss does not necessarily translate to unbiased improvements in decision-focused applications.

To address the research gap, this article introduces a supervised data purification framework designed to extract optimal high-quality datasets that minimize the cost of data-driven decision-making. We propose a two-stage valuation and filtering method to support power forecasting and day-ahead dispatch, aiming to refine datasets under uncertainty. Our contributions are succinctly categorized into three primary areas as follows.

1) We implement data purification within two-stage unit commitment (UC), achieving significant reductions in power dispatch costs by extracting high-quality datasets.
2) We construct a supervised data valuation approach tailored to day-ahead dispatch prediction needs. Leveraging deep reinforcement learning (RL), this framework dynamically updates strategies for identifying data quality.
3) We propose an iterative filtering mechanism that ensures precise quantification of essential data scales, guided by power dispatch outcomes.

II. High-Quality Data for Improved Decision-Making

In contrast to traditional methods that primarily rely on market mechanisms, statistical characteristics, or predictive models, our proposed data purification approach focuses on analyzing the uneven contributions of data to decision-making. Certain data features may hold greater decision value in specific contexts, while others may introduce noise or irrelevant information. For example, the same dataset may exhibit significantly different value distributions when applied to power system scheduling under normal versus extreme weather conditions. Our method aims to uncover the differentiated value of data for specific tasks by dynamically evaluating its actual contribution. This enables the prioritization of high-value data, supervised enhancing both the interpretability and optimization of decision-making processes.

In a general data-driven stochastic optimization framework, let ξ represent the uncertainty in the decision model J, such as renewable energy uncertainty in power dispatch problems. By leveraging the power of machine learning to capture nonlinear data patterns, ξ can be estimated using predictive models built on the dataset $\{(\mathbf{x}_i, \mathbf{y}_i)\}$ collected via sensors. Here, $\mathbf{x}_i$ denotes input data with multisourced features, such as multilevel meteorological information, while $\mathbf{y}_i$ represents the target output required for precise decision-making. Given that individual data samples are aggregated to construct the dataset used in the training process, the criterion for assessing data validity within the data purification framework is whether the high-quality dataset—refined through data valuation and filtering from the original dataset—effectively reduces the overall system decision-making cost.

Considering the two-stage modeling framework of power dispatch against renewable uncertainty, comprising day-ahead prediction and real-time decision-making [20], we propose a

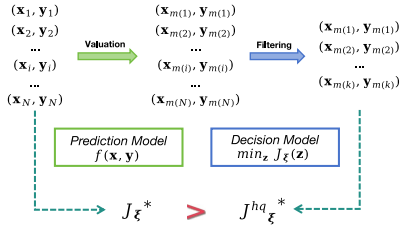

Fig. 1. Two-stage data purification.

two-stage data purification strategy that integrates data valuation and filtering, as depicted in Fig. 1. In the first phase, we design a power forecasting modeling tailored to decision-making needs and assess data quality by iteratively screening and evaluating their efficacy based on deep RL. This process reorganizes the data into a ranked sequence, denoted as m, based on its estimated quality. In the subsequent decision-making phase, we introduce an iterative filtering mechanism to determine the optimal high-quality dataset scale k based on the previous ranking m. In summary, the valuation phase differentiates the quality of various data, while the subsequent filtering phase refines the dataset with decision-focused adjustments. This interaction considers not only the influence of day-ahead forecasting on real-time dispatch, but also the practical computational efficiency of the process. Detailed explanations of these mechanisms are provided in Section IV. Ultimately, this two-stage method improves performance ($J_{\xi}^{\mathrm{hq}*}$) using the high-quality data subset, compared to the outcome without data purification (J_{ξ}^{*}).

III. Power Dispatch Formulation

This section presents a detailed formulation of the two-stage UC problem, a typical power dispatch challenge that considers unit startup decisions and the balancing of power demand and supply [21]. In day-ahead operations, we recognize wind power as a significant and dominant source of uncertainty, while other sources, such as load variations, are relatively minor by comparison [22].

A. Wind Power Forecasting

In this study, we focus on predicting wind power output for the forthcoming $T = 24$ h to support day-ahead decision-making in UC. Our approach enriches the dataset by combining time series data from the previous $kT = 48$ h (given $k = 2$) with hourly meteorological data, incorporating M weather features, such as *temperature*, *humidity*, *pressure*, *wind speed and direction*, *precipitation*, and *photosynthetically active radiation*. This fusion of historical and meteorological data transforms the forecasting task into a $kT + M$-dimensional problem, aimed at predicting power output for the next T h, formalized as $f : \mathbb{R}^{kT+M} \to \mathbb{R}^{T}$, correspond to $\mathbf{x}$ and $\mathbf{y}$ in the data structure definition in II.

For our predictive model, we selected the light gradient boosting machine (LightGBM), a decision tree-based ensemble learning technique known for its ability to handle high-dimensional data efficiently [23]. LightGBM partitions features using histogram-based algorithms and applies gradient boosting techniques to sequentially build a series of weak learners (decision trees) that iteratively improve model performance. The LightGBM model combines the outputs of these decision trees into a strong predictor, which can be expressed as follows:

$$\hat{\mathbf{y}} = \sum_{i=1}^{Ts} g^{(i)}(\mathbf{x}) \tag{1}$$

where Ts is the total number of trees, $g^{(i)}$ represents the output of the ith tree, and $\hat{\mathbf{y}}$ denotes the cumulative prediction for the target variable.

This approach makes it particularly effective for large-scale regression tasks involving multisource features, as it excels in capturing patterns within time-series data while maintaining high training speed and low memory usage, making it ideal for our wind power forecasting task [24]. The training objective of the LightGBM model is to minimize the following function under N training samples:

$$\arg\min_{g^{(i)}} \sum_{i=1}^{N} \mathcal{L}\left(\mathbf{y}_i, \hat{\mathbf{y}}_i^{(i-1)} + g^{(i)}\left(\mathbf{x}_i\right)\right) + \Omega\left(g^{(i)}\right) \tag{2}$$

where $\mathcal{L}$ represents the prediction loss function (e.g., MSE), $g^{(i)}$ is the output of the ith tree in the ensemble, and $\Omega(g^{(i)})$ is a regularization term to prevent overfitting by penalizing overly complex trees. $\hat{\mathbf{y}}_i^{(i-1)}$ denotes the cumulative predictions from the first $i-1$ trees, and the ith tree $g^{(i)}$ is trained to minimize the residuals $\mathbf{y}_i - \hat{\mathbf{y}}_i^{(i-1)}$ from the predictions of the existing ensemble, ensuring each tree contributes incremental improvements. By iteratively refining the model predictions and penalizing complexity, LightGBM effectively captures nonlinear relationships in data, making it highly suitable for our wind power forecasting task.

B. Unit Commitment Problem Design

Following wind power forecasting, the detailed formulation of the UC problem within a two-stage stochastic optimization framework, incorporating wind power uncertainty (ξ), is described as follows:

$$\begin{aligned}\min J_{\xi} = & \sum_{t=1}^{T}\sum_{i\in G}\left\{\alpha_{i,t}C_i^U + \beta_{i,t}C_i^D + \mathbb{E}\left[F_{c,i}\left(P_{i,t}^{G}(\xi)\right)\right]\right\} \\ & + \sum_{t=1}^{T}\sum_{i\in W} C^P \cdot \mathbb{E}\left[P_{i,t}^{W+}(\xi) + P_{i,t}^{W-}(\xi)\right]\end{aligned} \tag{3}$$

s.t.

$$-v_{i,t-1} + v_{i,t} - \alpha_{i,t} \le 0 \quad (\forall i \in G \quad t = 1, 2, \ldots, T) \tag{4}$$

$$v_{i,t-1} - v_{i,t} - \beta_{i,t} \le 0 \quad (\forall i \in G \quad t = 1, 2, \ldots, T) \tag{5}$$

$$\begin{aligned}&(v_{i,t+1} - v_{i,t}) + (v_{i,t+\sigma} - v_{i,t+\sigma+1}) \le 1 \\ &\quad (\forall i \in G \quad \forall \sigma \in [1, \cdots T_i^{\mathrm{ON}} - 1], t = 1, 2, \ldots, T)\end{aligned} \tag{6}$$

$$\begin{aligned}&(v_{i,t} - v_{i,t+1}) + (v_{i,t+\sigma+1} - v_{i,t+\sigma}) \le 1 \\ &\quad (\forall i \in G \quad \forall \sigma \in [1, \cdots T_i^{\mathrm{OFF}} - 1], t = 1, 2, \ldots, T)\end{aligned} \tag{7}$$

$$v_{i,t}P^G_{i,\min} \le P^G_{i,t}(\xi) \le v_{i,t}P^G_{i,\max} \quad (\forall i \in G \quad t = 1,2,\ldots,T) \tag{8}$$

$$P^G_{i,t+1}(\xi) - P^G_{i,t}(\xi) \le R^{\text{up}}_i v_{i,t} + P^G_{i,\text{up}}(1 - v_{i,t}) \quad (\forall i \in G \quad t = 1,2,\ldots,T) \tag{9}$$

$$P^G_{i,t}(\xi) - P^G_{i,t+1}(\xi) \le R^{\text{down}}_i v_{i,t+1} + P^G_{i,\text{down}}(1 - v_{i,t+1}) \quad (\forall i \in G \quad t = 1,2,\ldots,T) \tag{10}$$

$$P^G_{i,t}(\xi) + P^{AW}_{i,t}(\xi) = P^L_{i,t} + \sum_{j \in \phi_i} P_{ij,t}(\xi) \quad (\forall i \in U \quad t = 1,2,\ldots,T) \tag{11}$$

$$P_{ij,\min} \le P_{ij,t}(\xi) \le P_{ij,\max} \quad (\forall i \in U \quad t = 1,2,\ldots,T) \tag{12}$$

$$0 \le P^{AW}_{i,t}(\xi) \le P^W_{i,t}(\xi) \quad (\forall i \in W \quad t = 1,2,\ldots,T) \tag{13}$$

$$P^{AW}_{i,t}(\xi) - P^{FW}_{i,t} = P^{W+}_{i,t}(\xi) - P^{W-}_{i,t}(\xi) \tag{14}$$

$$0 \le P^{W+}_{i,t}(\xi), 0 \le P^{W-}_{i,t}(\xi) \quad (\forall i \in W \quad t = 1,2,\ldots,T)$$

$$S_{i,t}(\xi) = \max\{P^{FW}_{i,t} - P^W_{i,t}(\xi), 0\} \quad (\forall i \in W \quad t = 1,2,\ldots,T) \tag{15}$$

$$\sum_{i \in U} \eta P^L_{i,t} + \sum_{i \in W} S_{i,t}(\xi) \le \sum_{i \in G} \left(v_{i,t}P^G_{i,\max} - P^G_{i,t}(\xi)\right) \quad (t = 1,2,\ldots,T) \tag{16}$$

$$v_{i,t}, \alpha_{i,t}, \beta_{i,t} \in \{0,1\} \quad (\forall i \in G \quad t = 1,2,\ldots,T). \tag{17}$$

The definitions of parameters and variables are provided in the Nomenclature section. The objective function of the optimization problem (3) is structured in two stages. In the day-ahead scheduling phase, since generators cannot be temporarily switched ON or OFF in real time T due to grid safety considerations, the operational statuses of all generators are predetermined based on forecasts, incorporating the startup (C^U_i) and shutdown (C^D_i) costs of each unit. In the subsequent real-time dispatch phase, the expected cost is evaluated under the distribution of the stochastic variable ξ. For each unit i and time t, the operational generation $P^G_{i,t}$ incurs a fuel cost expressed as $F_{c,i}(P) = a_i P^2 + b_i P$. In addition, penalty costs is designed based on market mechanisms, with coefficients C^P, are imposed for misestimations of actual power production. These penalties are modeled using positive auxiliary variables $P^{W+}_{i,t}$ and $P^{W-}_{i,t}$. Under the two-stage formulation, variables related to wind power uncertainty in the second stage are represented as stochastic functions of the random vector ξ, while other variables based on forecasted results are determined in the first stage.

Operational constraints are laid out to ensure the system's integrity and efficiency: startup (4) and shutdown (5) protocols, minimum operational times (6), (7), alongside generation capacity (8), ramp-up (9), and ramp-down (10) flexibilities. System equilibrium is maintained through power balance (11) and transmission capacity (12) stipulations. Wind power integration is governed by output limits (13) and deviation constraints (14), with (15) quantifying potential energy deficits when wind generation falls short of commitments. Notably, the introduction of auxiliary variables can linearize the otherwise nonlinear constraint forms associated with maximum values in (15). Reserve requirements (16) for wind variability and load fluctuation, along with binary decision constraints (17) for initial planning stage, complete the framework designed to optimize UC in the face of wind power uncertainty.

C. Scenarios Generation

As a numerical simulation method for handling the expectation terms in (3), we employ Monte Carlo scenario generation to effectively create diverse scenarios that capture potential variations of uncertain elements in stochastic optimization [20]. Specifically, we model the distribution of wind power output $P^W(\xi)$ as a multivariate normal distribution $N(\mu, \Sigma)$ for each time period T, where μ represents the forecasted wind power and Σ its volatility. The parameters μ and Σ are fitted using the empirical distribution observed within the validation set [25].

Monte Carlo simulations sample from the distribution to generate scenarios $s \in \{1, 2 \ldots N_s\}$, each representing a distinct realization of the UC problem and occurring with probability π_s. Consequently, based on (3), the final operational objective of the two-stage UC problem is expressed as

$$\sum_{t=1}^{T} \sum_{i \in G} \left[\alpha_{i,t} C^U_i + \beta_{i,t} C^D_i + \sum_{s=1}^{N_s} \pi_s F_{c,i}\left(P^G_{i,t,s}\right) \right] + \sum_{t=1}^{T} \sum_{i \in W} \sum_{s=1}^{N_s} \pi_s C^P \cdot \left(P^{W+}_{i,t,s} + P^{W-}_{i,t,s}\right) \tag{18}$$

where the variables coupled with uncertainty ξ are decomposed into multiple probabilistic scenarios indexed by s, with N_s denoting the total number of Monte Carlo samples. Each scenario generation corresponds to a deterministic realization of all variables under a specific probability ($P^G_{i,t,s}$, $P^{W+}_{i,t,s}$, and $P^{W-}_{i,t,s}$), transforming the original theoretical model into a solvable, scenario-based final formulation (18).

Accordingly, the stochastic constraints in (4)–(17) are also reformulated within the scenario generation framework, ensuring consistency across the UC problem.

IV. Data Purification Mechanism

This section comprehensively introduces the concept of data purification through three sections on valuation—formulation of supervised data valuation (A), transition to RL (B), and policy exploration (C)—and one section on cost-oriented data filtering (D). Stochastic error serves as a key reference for penalty costs in (18) and provides a computationally efficient proxy, avoiding the direct integration of complex and resource-intensive power system optimization. Consequently, data valuation leverages prediction metrics to align dataset quality with the decision-making requirements of day-ahead UC over T horizons.

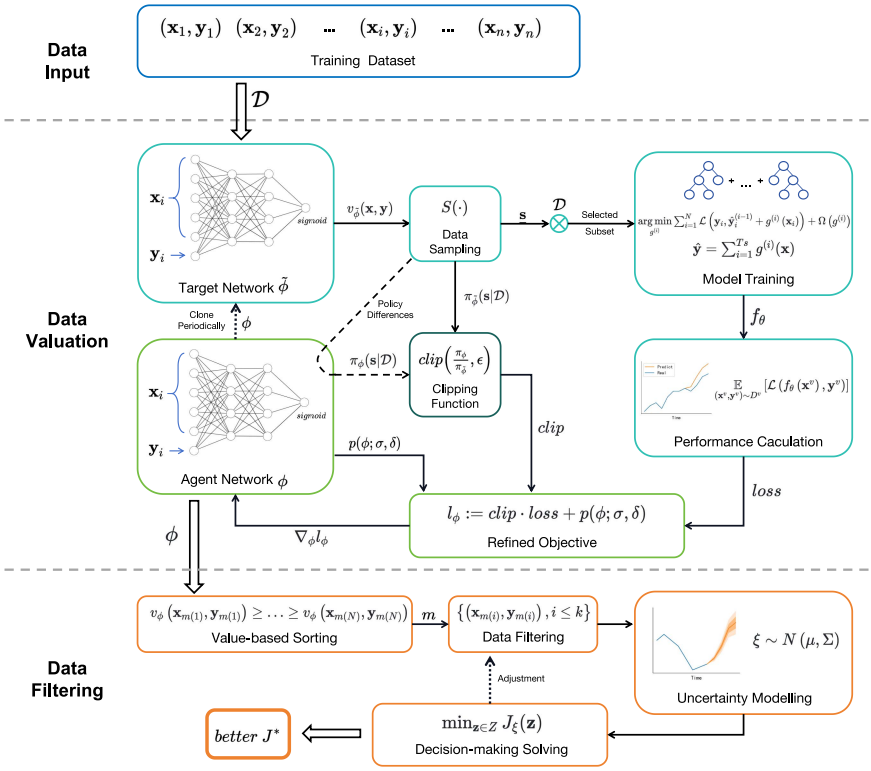

Fig. 2. Framework of DPID.

After presenting the theoretical supervised framework, we address key challenges, including nondifferentiability and convergence issues, by incorporating RL and refining policy strategies. Finally, leveraging the valuation results with high-quality data sequence, we determine the optimal data scale based on dispatch costs to implement decision-oriented filtering. This comprehensive approach, termed data purification for improved decision-making (DPID), encompasses the entire process from data input, valuation, to filtering, as illustrated in Fig. 2.

A. Supervised Data Valuation

The objective function is designed to extract high-quality data, focusing on optimizing predictive accuracy within the given dataset. To avoid conceptual ambiguity, we define "data value" as the quantified measure of data quality. We assess the dataset $\mathcal{D} := \{(\mathbf{x}_i, \mathbf{y}_i)\}$, where $i \in \{1 \ldots N\}$, to extract a high-quality subset conducive to predictive modeling. Validation of the model's performance employs a separate dataset $\mathcal{D}^v := \{(\mathbf{x}_i^v, \mathbf{y}_i^v)\}$. We denote the distributions of the training and validation datasets as D and D^v, respectively.

To validate the effectiveness of data valuation, we recalibrate the predictive model by retraining it using the selected high-quality data to observe its impact on prediction performance. For this purpose, we designed a data sampling function $s_\phi : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$, which identifies high-quality data points (1 for high-quality, 0 for low-quality) based on their features and corresponding labels. This function s_ϕ filters the original dataset $\mathcal{D}$, isolating a high-quality subset. To effectively handle high-dimensional features, such as numerical weather predictions and historical power sequences, the sampling function is implemented as a deep learning network with parameters ϕ. This design facilitates the precise identification of potential nonlinear and intricate mappings from features to value, which are critical for a comprehensive assessment of data quality.

The goal of supervised valuation is to train ϕ to evaluate the value of each data point in D, filtering a high-quality subset for constructing the prediction model f_θ, which minimizes the loss on the validation set D^v. The predefined prediction loss metric $\mathcal{L}$ aligns with the training objective in (2). The fundamental valuation process is formalized as follows:

$$\min_{\phi} \quad \mathop{\mathbb{E}}_{(\mathbf{x}^v, \mathbf{y}^v) \sim D^v} \left[\mathcal{L} \left(f_\theta \left(\mathbf{x}^v \right), \mathbf{y}^v \right) \right] \tag{19}$$

$$\text{s.t. } f_\theta = \arg\min_{f} \mathop{\mathbb{E}}_{(\mathbf{x}, \mathbf{y}) \sim D} \left[s_\phi \left(\mathbf{x}, \mathbf{y} \right) \mathcal{L}(f(\mathbf{x}), \mathbf{y}) \right] \tag{20}$$

where (19) minimizes the prediction loss on the validation set D^v, and (20) describes retraining the predictive model f_θ on a high-quality subset selected by the sampling function s_ϕ.

However, this fundamental formulation presents a challenge: The sampling function s_ϕ, which maps each data point's feature space ($\mathcal{X} \times \mathcal{Y}$) to a binary domain ($\{0,1\}$), introduces nondifferentiability in the iterative training process. This occurs because gradient information is truncated during discrete selection, hindering precise back-propagation and gradient-based updates. To address this issue, we integrate RL into the framework, enabling gradient information to flow through the sampling process and achieve unbiased training.

B. Reinforcement Learning Formulation

In this section, we reformulate the initial objective in (19) within an RL framework, utilizing policy gradient methods to

achieve unbiased gradient computation by taking the derivative of the selection probability in its log form. The selection function is decomposed into two coupled components: A deep learning-based valuation network $v_\phi : \mathcal{X} \times \mathcal{Y} \to [0, 1]$ and a subsequent binomial probability-based sampling module $S : [0, 1] \to \{0, 1\}$. This decomposition is expressed as

$$s_\phi = S \circ v_\phi. \tag{21}$$

By using continuous outputs from v_ϕ, gradient information can propagate within the network through backpropagation. The high-quality subset selection is then transformed into a stochastic action, where the output of v_ϕ represents the probability of a data point being high-quality. Data points with values closer to 1 are more likely to be selected, while those closer to 0 are more likely to be filtered out, which provides an interpretable explanation of data value.

With the data valuation extended to a continuous 0–1 space, the interpretation of stochastic selection in the RL environment is as follows: The dataset $\mathcal{D}$ is treated as the environment's state, the selection vector $\mathbf{s} := [s_\phi(\mathbf{x}_i, \mathbf{y}_i)]_{1:N}$ represents the action, and the objective in (19) corresponds to the negative reward for each iteration. In the environment $\mathcal{D}$, the selection vector generated based on the network's valuation is expressed probabilistically as

$$\pi_\phi(\mathbf{s}|\mathcal{D}) = \prod_{i=1}^{N} \left[v_\phi(\mathbf{x}_i, \mathbf{y}_i)^{s_i} \cdot (1 - v_\phi(\mathbf{x}_i, \mathbf{y}_i))^{1-s_i} \right] \tag{22}$$

where s_i is the ith component of $\mathbf{s}$. This reformulation allows the initial objective in (19) to be extended into a stochastic selection-based version

$$\mathop{\mathbb{E}}_{\substack{(\mathbf{x}^v, \mathbf{y}^v) \sim D^v \\ \mathbf{s} \sim \pi_\phi(\cdot|\mathcal{D})}} \left[\mathcal{L}\left(f_\theta(\mathbf{x}^v), \mathbf{y}^v \right) \right]. \tag{23}$$

We consider a single-step task with a static environment state $\mathcal{D}$. Using the policy gradient method [26], the gradient ∇_ϕ of the RL objective (23) into a product of the validation loss and the log of the action probability, expressed as

$$\mathop{\mathbb{E}}_{\substack{(\mathbf{x}^v, \mathbf{y}^v) \sim D^v \\ \mathbf{s} \sim \pi_\phi(\cdot|\mathcal{D})}} \left[\mathcal{L}\left(f_\theta(\mathbf{x}^v), \mathbf{y}^v \right) \cdot \nabla_\phi \log\left(\pi_\phi(\mathbf{s}|\mathcal{D}) \right) \right]. \tag{24}$$

This formulation theoretically ensures unbiasedness and, as shown by the probabilistic expression in (22), establishes that the right-hand side term in (24) is differentiable due to the continuous output of v_ϕ. Consequently, the RL objective in (23) can be trained using gradient descent or similar optimization algorithms. This process involves iterative subset sampling, where each subset $\mathbf{s} \times \{(\mathbf{x}_i, \mathbf{y}_i)\}_{1:N}$ is generated at each training epoch.

C. Policy Refinements

This section aims to improve the stability and convergence of the data valuation process within the RL framework by proposing refinements to the objective (23).

To address the challenge of balancing exploration and exploitation, which can hinder overall learning efficiency [27], and to enhance the utilization efficiency of training samples, we refine the objective (23) using an importance sampling method as follows:

$$\mathop{\mathbb{E}}_{\substack{(\mathbf{x}^v, \mathbf{y}^v) \sim D^v \\ \mathbf{s} \sim \pi_{\tilde{\phi}}(\cdot|\mathcal{D})}} \left[\frac{\pi_\phi(\mathbf{s}|\mathcal{D})}{\pi_{\tilde{\phi}}(\mathbf{s}|\mathcal{D})} \mathcal{L}\left(f_\theta(\mathbf{x}^v), \mathbf{y}^v \right) \right]. \tag{25}$$

Here, $\tilde{\phi}$ represents the parameters of the target network, which shares the same dimensionality as ϕ. In this refined objective, the training samples for each iteration are generated using $\tilde{\phi}$, which operates independently of the gradient-updated parameters ϕ. This decoupling increases the number of learning steps under a consistent valuation policy (off-policy), enhancing the robustness of stochastic policy updates. After every policy cycle of c iterations, $\tilde{\phi}$ is periodically synchronized with ϕ, preventing excessive divergence between the sampling network and the updated network. To ensure unbiasedness, compared to the previous objective (23), this refined objective introduces a probability ratio $\frac{\pi_\phi}{\pi_{\tilde{\phi}}}$ between the selection probabilities of the two networks, applied as a weighting factor to the prediction loss.

To further mitigate the variability introduced by differences between the two networks during training, we adopt a clipping mechanism for probability ratio in (25), inspired by proximal policy optimization [28], which constrains the range of effective gradient updates

$$\text{clip}\left(\frac{\pi_\phi}{\pi_{\tilde{\phi}}}, \epsilon \right) = \max\left(\min\left(1 - \epsilon, \frac{\pi_\phi}{\pi_{\tilde{\phi}}} \right), 1 + \epsilon \right) \tag{26}$$

where ϵ is a small parameter that restricts the effective selection to cases where the probabilities of the two networks are similar. When the deviation exceeds the threshold ϵ, the gradient becomes zero, as the terms involving ϕ in the objective (25) are replaced with either $1 - \epsilon$ and $1 + \epsilon$. Other cases, the gradient (∇_ϕ) of refined objective in (25) with clipped mechanism in (26), are similar in form to (24)

$$\mathop{\mathbb{E}}_{\substack{(\mathbf{x}^v, \mathbf{y}^v) \sim D^v \\ \mathbf{s} \sim \pi_{\tilde{\phi}}(\cdot|\mathcal{D})}} \left[\frac{\pi_\phi}{\pi_{\tilde{\phi}}} \mathcal{L}\left(f_\theta(\mathbf{x}^v), \mathbf{y}^v \right) \cdot \nabla_\phi \log\left(\pi_\phi \right) \right]. \tag{27}$$

This mechanism enhances the robustness and efficiency of our data valuation methodology by ensuring stable updates while maintaining unbiasedness.

To prevent the valuation function v_ϕ from converging to trivial local optima—either assigning all outputs to 0 or 1—we introduce a two-sided penalty term. The specific expression is given as

$$\begin{aligned} p(\phi; \sigma, \delta) = {} & \sigma \cdot \max\left(\sum_{i=1}^{N} v_\phi(\mathbf{x}_i, \mathbf{y}_i) - N \cdot \delta, 0 \right) \\ & + \sigma \cdot \max\left(N - \sum_{i=1}^{N} v_\phi(\mathbf{x}_i, \mathbf{y}_i) - N \cdot \delta, 0 \right) \end{aligned} \tag{28}$$

where δ is a small positive parameter that acts as a threshold to encourage balanced valuation outputs, and σ is the penalty

coefficient. This design ensures that the model avoids degenerate solutions where all data points are deemed either entirely unimportant or uniformly significant, thereby mitigating the influence of noise and promoting meaningful data differentiation.

In summary, combined with importance sampling in (25), clipping mechanism (26) and suboptimal penalty in (28), the final objective of RL data valuation in term of deep-learning network parameters ϕ is

$$\underset{\substack{(\mathbf{x}^v,\mathbf{y}^v)\sim D^v \\ \mathbf{s}\sim\pi_{\tilde{\phi}}(\cdot|\mathcal{D})}}{\mathbb{E}}\left[\text{clip}\left(\frac{\pi_\phi}{\pi_{\tilde{\phi}}},\epsilon\right)\mathcal{L}\left(f_\theta\left(\mathbf{x}^v\right),\mathbf{y}^v\right)\right]+p(\phi;\sigma,\delta). \tag{29}$$

Combined with the constraint in (20) and the refined policy gradient expression in (27) for the expectation term, the valuation phase is trained in an unbiased, efficient, and robust manner.

D. Cost-Oriented Data Filtering

Following the ranking of the converged value outputs, as illustrated in Fig. 1, the implementation of a strategic filtering process is critical for determining the optimal dataset size for UC decision-making. Insufficient dataset sizes risk omitting critical information during model training, thereby compromising performance and the efficiency of the UC process. Conversely, excessively large datasets can dilute the efficiency and cost-effectiveness of the data valuation and filtering process, undermining the benefits of precise data governance.

To align data value with the practical cost implications of the UC objective, we introduce a data filtering mechanism to finalize the decision-focused data purification. This mechanism assumes that the relationship between data volume—from the upper bound u to lower bound l —and the objective function is approximately convex. This assumption implies that initially removing lower-quality data points reduces costs by improving decision-making efficiency. However, excessive exclusion may result in the loss of essential data, which can inversely increase costs. This convexity assumption, grounded in data utilization theories from machine learning, is empirically validated in our case studies, demonstrating its applicability to the UC problem.

The data filtering algorithm, employed by golden-section based on the value sorting as m, is demonstrated in Fig. 2, proceeds as follows.

1) *Step 1:* Initialize upper-bound u and lower-bound l with $u > l$.
2) *Step 2:* Compute J_u and J_l in (18) using dataset $\{(\mathbf{x},\mathbf{y})_{m[1:u]}\}$ and $\{(\mathbf{x},\mathbf{y})_{m[1:l]}\}$, respectively, including both predictive model training (detailed in Section II-I-A) and optimization problem solving (detailed in Sections III-B and III-C).
3) *Step 3:* If $u - l <$ mini_unit, *return* $\min(J_u, J_l)$.
4) *Step 4:* $u' \leftarrow u - \rho(u-l)$ and $l' \leftarrow l + \rho(u-l)$.
5) *Step 5:* Compute $J_{u'}$ and $J_{l'}$ in (18) using dataset $\{(\mathbf{x},\mathbf{y})_{m[1:u']}\}$ and $\{(\mathbf{x},\mathbf{y})_{m[1:l']}\}$, respectively.
6) *Step 6:* If $J_{u'} > J_{l'}$, $u \leftarrow u'$; else, $l \leftarrow l'$.Go back to *Step 3*.

Here, mini_unit represents the dataset's minimum segmentation scale, and parameter $\rho \in (0, 0.5)$ controls the convergence rate. $\{(\mathbf{x},\mathbf{y})_{m[1:k]}\}$ with $1 \le k \le N$ represents the selection of the top k highest value data points. Ultimately, this iterative approach narrows the upper and lower bounds to converge on the optimal high-quality dataset size, minimizing the cost of the UC problem.

At this stage, the two-stage data purification process, consisting of valuation and filtering, is complete. The detailed workflow is illustrated in Fig. 2, and the pseudocode for our DPID is presented in Algorithm 1.

Algorithm 1: Data Purification for Improved Decision.

Input: training dataset $\mathcal{D} := \{(\mathbf{x}_i, \mathbf{y}_i)\}$, validation dataset $\mathcal{D}^v := \{(\mathbf{x}_i^v, \mathbf{y}_i^v)\}$, training iteration M, prediction loss $\mathcal{L}$, learning rate α, penalty factor σ, value threshold δ, clipping parameter ϵ, policy update cycle c
Output: data value $v_\phi(\mathcal{D})$, improved Objective J^*
Initialize: parameters ϕ, $\tilde{\phi}$

for $k \leftarrow 1$ **to** M **do**
  if $c \mid k$ **then**
    $\tilde{\phi} \leftarrow \phi$
  end if
  for $i \leftarrow 1$ **to** N **do**
    $s_i \leftarrow S(v_\phi(\mathbf{x}_i, \mathbf{y}_i))$
    $\mathbf{s} \leftarrow [s_i]_{1:N}$
  end for
  train the predictive model:
    $f_\theta = \arg\min \quad \mathbf{s} \times \mathcal{L}(f(\mathbf{x}), \mathbf{y})$
  update the valuation parameter:
    if $\frac{\pi_\phi(\mathbf{s}|\mathcal{D})}{\pi_{\tilde{\phi}}(\mathbf{s}|\mathcal{D})} \in [1-\epsilon, 1+\epsilon]$ **then**
      $\hat{l}_\phi \leftarrow \frac{\pi_\phi(\mathbf{s}|\mathcal{D})}{\pi_{\tilde{\phi}}(\mathbf{s}|\mathcal{D})} \cdot \mathbb{E}[\mathcal{L}(f_\theta(\mathbf{x}^v), \mathbf{y}^v)]$
      $\phi \leftarrow$
    $\phi + \alpha \cdot \hat{l}_\phi \cdot \nabla_\phi \log(\pi_\phi(\mathbf{s}|\mathcal{D})) + \alpha \cdot \nabla_\phi p(\phi; \sigma, \delta)$
    else then
      $\phi \leftarrow \phi + \alpha \cdot \nabla_\phi p(\phi; \sigma, \delta)$
    end if
  if convergence **then**
    break
  end if
end for
return $v_\phi(\mathcal{D})$
$m \leftarrow sort(v_\phi(\mathcal{D}))$
Execute *Step 1-Step 6*
return J^*

V. Numerical Results

In this study, we utilize hourly wind power output data from Sichuan Province (2017–2018) [29], dividing the dataset into training and validation sets with an 8:2 chronological split. To construct UC problems, we adopt the IEEE 30-bus and 118-bus systems, ensuring that node and network parameters align with established standards in the literature [30]. In the IEEE 30-bus system, three wind farm are connected to nodes 10, 15, and 17,

TABLE I
DATA FILTERING ADJUSTMENT BASED ON UC COST

Iteration	u (%)	l (%)	J_u (10^5)	J_l (10^5)
1	100	50	1.0685	1.0492
2	80	50	1.0540	1.0492
3	70	50	1.0501	1.0492
4	70	60	1.0501	1.0461
5	65	60	1.0487	**1.0461**

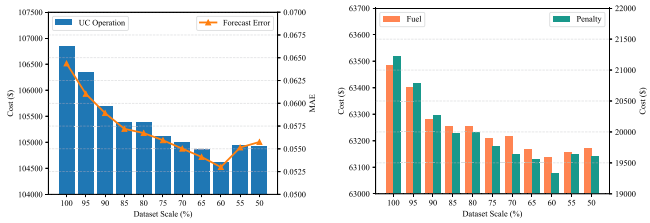

Fig. 3. Operation cost and forecast error (left) and fuel and penalty cost (right) on different scale of valued dataset.

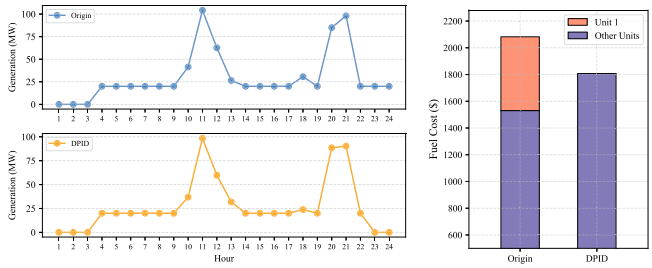

Fig. 4. Power generation of unit 1 (left) and cost on other units at the 24th h.

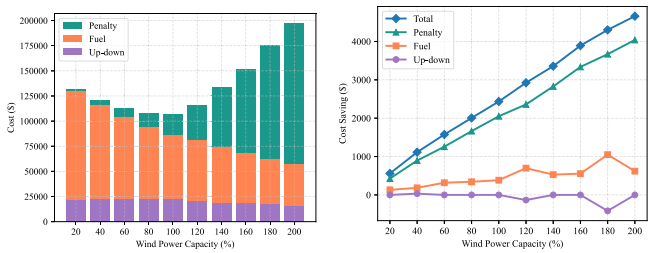

Fig. 5. Cost components (left) and cost saving under data purification (right) on various wind power capacities.

with a penalty coefficient of 100/MWh. The measured datasets from these three wind arrays are randomly selected and tested for the forecasting task, as referenced in [11]. In addition, the load curves at each node are sampled from actual load data sourced from Texas, reflecting realistic operational contexts [31]. For critical algorithm parameters, we set $\mathcal{L}$ as MAE, α as *Adam*, $\sigma = 1e6, \delta = 0.95, \epsilon = 0.2$, and $c = 3$. All experiments were repeated multiple times, and the results were averaged.

A. Case Study of the IEEE 30-Bus System

1) High-Quality Data for Operation Cost Reduction: In the experiments conducted on the IEEE-30 system, by setting the mini-unit to mini_unit $= 5\%$, $u = 100\%$, $l = 50\%$, and $\rho = 0.4$, we successfully achieved optimal data filtering (60% high-value data) within five iterations, following valuation convergence. The results are documented in Table I.

To further investigate the relationship between high-quality data scale and its impact on prediction and decision-making, we present the wind power forecast error and UC operation cost under the 5% segment setting in Fig. 3 (left). The consistent results validate our convex hypothesis in data filtering, demonstrating that removing low-value data points can enhance forecast accuracy, reducing prediction errors by up to 13.2%. However, when the remaining dataset size falls below 60%, where the current lowest value data points are no longer as insignificant as those in the original dataset, critical features may be discarded. This leads to a reverse trend, with both prediction errors and decision-making costs increasing.

By decomposing the dispatch cost into fuel cost (F) and penalty cost (P) as defined in (3), we observed consistent trends, as illustrated in Fig. 3 (right). The penalty cost is directly influenced by the accuracy of wind power output distribution estimation. As the precision of the data-driven model improves, uncertainty estimation becomes more aligned with real-world scenarios, leading to a reduction in penalty costs. Correspondingly, the difference in fuel cost arises from adjustments in energy reserves to accommodate wind power uncertainty. The removal of 40% low-value data significantly enhances modeling accuracy, enabling the system to require fewer reserves and increasing operational flexibility for low-cost generators.

To illustrate how data purification reduces the UC cost, we analyze the intermediate variable of Unit 1's generation output under abundant wind resource conditions, as shown in Fig. 4 (left). Unit 1, characterized by the highest capacity ($P^G_{\max}$) and generation cost coefficients (F_c), incurs significant operating expenses, necessitating careful evaluation to justify its dispatch. Before hour 22, the generation patterns in the DPID and original cases are comparable. However, after hour 22, notable differences arise as DPID's advanced uncertainty estimation reduces reliance on high-cost storage (S). This enables the system to confidently deactivate the expensive Unit 1, leveraging the spare capacity of other units. Although the other units increase their output to maintain power balance, the overall fuel cost is lower compared to the original case, which keeps Unit 1 online as a precaution against potential fluctuations. This comparison, illustrated in Fig. 4 (right), highlights DPID's superior efficiency in optimizing power dispatch by precisely controlling uncertainty and shifting generation to lower-cost units.

2) Sensitivity and Transferability Analysis: We investigate the robustness of the proposed algorithm under varying wind power capacities by constructing scenarios with capacities ranging from 20% to 200% in 20% intervals. Without data purification, Fig. 5 (left) shows that increasing wind capacity reduces fuel costs but raises penalty costs and operational risks, resulting in a total cost curve that initially decreases and then rises. In Fig. 5 (right), data purification (discarding 40% of low-value data) achieves a linear increase in cost savings as wind capacity expands, with a 2.3% reduction observed at baseline capacity (100%). These savings are primarily driven by reductions in penalty and fuel costs, resulting from improved power estimation accuracy and lower operational risks, highlighting the

TABLE II
EFFECTS OF PARTIALLY TRAINED DATA PURIFICATION ON UNTRAINED AND WHOLE DATASETS (USD)

	Remove low	Remove high	Difference
Untrained dataset	+5213.27	−9152.84	14366.11
Whole dataset	+2503.58	−8324.66	10828.24

decision-making benefits of data purification in wind-abundant scenarios.

Achieving the generalization of data purification requires addressing the potential risk of overfitting, particularly in ensuring that a valuation strategy trained on forecast tasks can effectively adapt to unseen data. In Section IV, we addressed this by designing a valuation framework emphasizing transferability. Instead of a static or parameterized pricing strategy, we employed a learnable neural network to ensure possible outputs for structured data. In addition, RL-based refinements were introduced to enhance sampling robustness in dynamic environments.

To evaluate generalization, we partitioned the dataset into a 5:3:2 split, with 37.5% of the original training data held out as an untrained subset. Training was limited to the top 50% of the dataset, and the valuation and filtering performance were tested on both the untrained subset and the original training set. Results shown in Table II, assessed using unpurified UC costs, demonstrate that the partially trained model effectively reduced decision-making costs on the untrained subset by correctly identifying and retaining (or removing) high-value data, even for previously unseen data. On the whole dataset, it achieved operational cost reductions comparable to the original task. These results demonstrate the transferability of the proposed data purification approach, generalizing value patterns across datasets and ensuring accurate valuation for unseen distributions.

3) Comparison of Data Purification Methods: In this chapter, we compare several baseline models, including the following. 1) Outlier detection using the Z-Score, a general method commonly applied in data purification tasks [10]; 2) Shannon EN, widely used in power systems [11], computed using a Gaussian kernel Parzen window to estimate the empirical distribution of the samples, defined as

$$H(\mathcal{D}) = -\sum_i p_i \log_2(p_i) \tag{30}$$

where p_i represents the estimated probability density of each sample; and 3) Data Shapley value (SV) [14], a cooperative game-theoretic method that quantifies the marginal contribution of each data point to the model's performance. The SV Φ for a data point $\mathcal{D}_i$ in dataset $\mathcal{D}$ is given by

$$\Phi(\mathcal{D}_i) = \sum_{S \subseteq \mathcal{D} \backslash \{\mathcal{D}_i\}} \frac{|S|!(N-|S|-1)!}{N!} \left[\mathcal{V}(S \cup \mathcal{D}_i) - \mathcal{V}(S)\right] \tag{31}$$

where $\mathcal{V}$ is the mapping from the training dataset to the negative prediction loss. Both outlier and EN methods are theoretically designed for 1-D arrays, so we sum all features of the dataset to obtain a single value per data point with multiple features. For

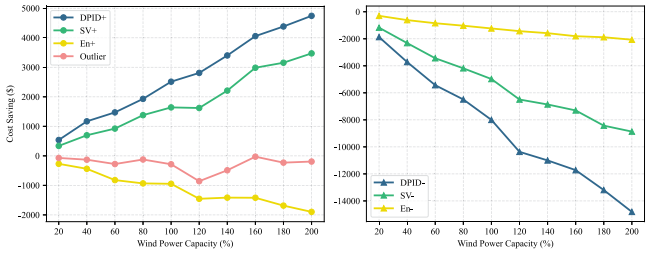

Fig. 6. Comparison of data purification methods on UC cost saving.

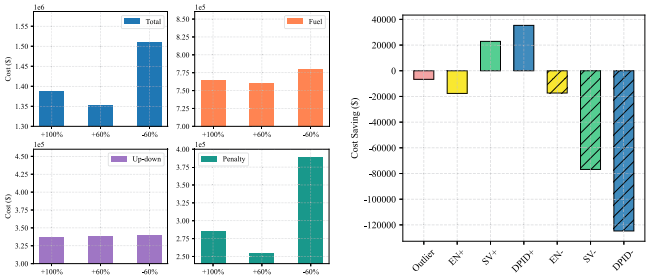

Fig. 7. Costs components on IEEE-118 (left) and data purification methods comparison (right).

EN, a leave-one-out approach is used to evaluate the marginal contribution of each data point.

In addition to outlier removal for only low-value data, with other methods, we use "+" to signify retaining 60% of high-value data, while "−" represents retaining 60% of low-value data. The cost savings of different methods compared to the 'no purification' baseline are shown in Fig. 6. Unsupervised methods, outlier and EN, exhibit lower valuation efficiency under large-scale feature data, struggling to distinguish data contributions to decision-making. In contrast, SV and DPID can accurately assess the contribution of different data points to the model, with the "+" and "−" labels highlighting expected differences. Notably, our proposed DPID outperforms SV in decision-making. The relative economic advantage of DPID is 1.60 times that of SV at 20% capacity, 1.53 times at 100% capacity, and 1.47 times at 200% capacity in UC. Furthermore, the damage and additional cost from removing high-value points in DPID are at least 1.60 times those in SV at these three points. These results not only demonstrate the superiority of our algorithm, but also emphasize the positive impact of more precise data value assessment on power dispatch under uncertainty.

B. Case Study of the IEEE 118-Bus System

To validate the proposed data purification approach in large-scale power systems, we conduct a study using the IEEE 118-bus system. Twelve wind farms are connected to nodes 11, 18, 24, 29, 54, 55, 62, 65, 83, 89, 110, and 115, with curves sampled from the validation set. We compare three strategies: retaining the entire dataset (+100%), utilizing 60% high-quality data (+60%), and employing 60% low-quality data (−60%), as depicted in Fig. 7 (left). The results align with the trends observed in the 30-bus system, but with greater absolute cost savings, attributable to the increased uncertainty scale from larger wind farm capacities. In addition, the heightened penalty costs from wind power

fluctuations in large-scale systems underscore the critical need for precise data purification to enable effective wind power prediction and control.

For further analysis in the IEEE 118-bus system, we conducted a comparative experiment where "+" retains high-value data and "−" retains low-value data, as shown in Fig. 7. Outlier-based methods had minimal impact, failing to distinguish the contributions of individual data points. EN, although commonly used in the energy sector, struggled to capture feature interactions in complex large-scale systems, showing limited effectiveness. In contrast, SV and our proposed DPID aligned with expected valuation trends, with DPID outperforming Shapley by 1.54 times in "+" actions and 1.62 times in "−" actions, slightly exceeding results from the IEEE 30-bus system. These results demonstrate DPID's superior performance in large-scale operations and its critical role in enabling significant economic savings in power systems.

VI. Conclusion

In this article, we propose a two-stage data purification method to enhance UC operations under wind power uncertainty. In the first stage, data quality is assessed using RL-based valuation, incorporating refined policy strategies tailored to forecasting requirements, which generates a prioritized reference to guide the decision-focused filtering in the second stage. In case studies on IEEE-30 and IEEE-118 systems, the proposed DPID method outperformed existing approaches in accurately identifying high-quality data and delivering economic benefits across varying uncertainty levels. This method offers a conceptual breakthrough for academia and industry by purifying high-quality datasets for targeted decision-making, providing essential methodological guidance for optimizing decisions in Big Data-driven energy systems.

Looking ahead, two essential avenues for further exploration emerge. First, exploring the possibility of end-to-end approaches that integrate valuation and filtering more directly. Second, addressing potential overfitting issues in the data purification process is essential for enhancing generalization and ensuring robust performance on larger unseen samples and dynamic domains.

References

[1] J. Wang et al., "Data sharing in energy systems," *Adv. Appl. Energy*, vol. 10, 2023, Art. no. 100132.

[2] X. Gong, X. Wang, and B. Cao, "On data-driven modeling and control in modern power grids stability: Survey and perspective," *Appl. Energy*, vol. 350, 2023, Art. no. 121740.

[3] S. Impram, S. V. Nese, and B. Oral, "Challenges of renewable energy penetration on power system flexibility: A survey," *Energy Strategy Rev.*, vol. 31, 2020, Art. no. 100539.

[4] A. Jain et al., "Overview and importance of data quality for machine learning tasks," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2020, pp. 3561–3562.

[5] I. Triguero, D. García-Gil, J. Maillo, J. Luengo, S. García, and F. Herrera, "Transforming Big Data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data," *Wiley Interdiscipl. Rev.: Data Mining Knowl. Discov.*, vol. 9, no. 2, 2019, Art. no. e1289.

[6] A. Agarwal, M. Dahleh, and T. Sarkar, "A marketplace for data: An algorithmic solution," in *Proc ACM Conf. Econ. Computation*, 2019, pp. 701–726.

[7] C. Goncalves, P. Pinson, and R. J. Bessa, "Towards data markets in renewable energy forecasting," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 533–542, Jan. 2021.

[8] B. Wang, Q. Guo, T. Yang, L. Xu, and H. Sun, "Data valuation for decision-making with uncertainty in energy transactions: A case of the two-settlement market system," *Appl. Energy*, vol. 288, 2021, Art. no. 116643.

[9] Z. Sun, L. Von Krannichfeldt, and Y. Wang, "Trading and valuation of day-ahead load forecasts in an ensemble model," *IEEE Trans. Ind. Appl.*, vol. 59, no. 3, pp. 2686–2695, May/Jun. 2023.

[10] A. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, "A review on outlier/anomaly detection in time series data," *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1–33, 2021.

[11] M. Yu et al., "Pricing information in smart grids: A quality-based data valuation paradigm," *IEEE Trans. Smart Grid*, vol. 13, no. 5, pp. 3735–3747, Sep. 2022.

[12] R. Mieth, J. M. Morales, and H. V. Poor, "Data valuation from data-driven optimization," *IEEE Trans. Control Netw. Syst.*, early access, Jul. 19, 2024, doi: 10.1109/TCNS.2024.3431415.

[13] W. Liang et al., "Advances, challenges and opportunities in creating data for trustworthy ai," *Nature Mach. Intell.*, vol. 4, no. 8, pp. 669–677, 2022.

[14] A. Ghorbani and J. Zou, "Data shapley: Equitable valuation of data for machine learning," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2019, pp. 2242–2251.

[15] R. Jia et al., "Efficient task-specific data valuation for nearest neighbor algorithms," in *Proc. VLDB Endowment*, vol. 12, no. 11, pp. 1610–1623, 2019.

[16] A. Ghorbani, M. Kim, and J. Zou, "A distributional framework for data valuation," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2020, pp. 3535–3544.

[17] Y. Wang, J. Wang, F. Gao, and J. Song, "Unveiling value patterns via deep reinforcement learning in heterogeneous data analytics," *Patterns*, vol. 5, no. 5, 2024, Art. no. 100965.

[18] Y. Wang and J. Song, "Dissecting renewable uncertainty via deconstructive analysis-based data valuation," *IEEE Trans. Ind. Appl.*, vol. 61, no. 1, pp. 1626–1635, Jan./Feb. 2024.

[19] J. Zhang, Y. Wang, and G. Hug, "Cost-oriented load forecasting," *Electric Power Syst. Res.*, vol. 205, 2022, Art. no. 107723.

[20] Q. P. Zheng, J. Wang, and A. L. Liu, "Stochastic optimization for unit commitment—A review," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1913–1924, Jul. 2015.

[21] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power Generation, Operation, and Control*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2013.

[22] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. Berlin, Germany: Springer Science & Business Media, 2011.

[23] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, vol. 30, pp. 3149–3157.

[24] X. Sun, M. Liu, and Z. Sima, "A novel cryptocurrency price trend forecasting model based on lightgbm," *Finance Res. Lett.*, vol. 32, 2020, Art. no. 101084.

[25] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 206–215, Feb. 2012.

[26] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT press, 2018.

[27] T. Zhao, H. Hachiya, G. Niu, and M. Sugiyama, "Analysis and improvement of policy gradient estimation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2011, vol. 24, pp. 262–270.

[28] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.

[29] X. Lu, M. B. McElroy, W. Peng, S. Liu, C. P. Nielsen, and H. Wang, "Challenges faced by China compared with the us in developing wind power," *Nature Energy*, vol. 1, no. 6, pp. 1–6, 2016.

[30] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Application of large-scale synthetic power system models for energy economic studies," 2017.

[31] J. Wang, J. Qin, H. Zhong, R. Rajagopal, Q. Xia, and C. Kang, "Reliability value of distributed solar storage systems amidst rare weather events," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4476–4486, Jul. 2019.