Decision-Informed Prediction Intervals of Renewable Energy Generation for Robust OPF

Zhiqiang He, Can Wan, Kaiming Zhang, Zhuo Xu
College of Electrical Engineering, Zhejiang University, Hangzhou, China
Emails: zhiqiang_ee@zju.edu.cn, canwan@zju.edu.cn, zhangkaiming@zju.edu.cn, zhuoxu@zju.edu.cn

***Abstract*—The large-scale integration of renewable energy generation introduces substantial uncertainty to power system operations. Most probabilistic forecasting approaches aim to improve forecasting accuracy but overlook the influence of downstream decision-making processes. To address this issue, the paper proposes a decision-informed prediction interval (DIPI) method that integrates forecasting and robust optimal power flow (ROPF) through a bilevel optimization-based framework. The upper-level model constructs PIs by minimizing the system operational cost, while the lower-level ROPF is reformulated as a differentiable linear program embedded within a convex optimization layer. An improved difference of convex functions-based surrogate is devised to handle the non-differentiable coverage constraint, which is combined with the augmented Lagrangian method to enable end-to-end training of deep learning models. Numerical studies on the IEEE 14-bus and 118-bus systems demonstrate that the DIPI method produces reliable prediction intervals and significantly reduces operational costs, highlighting the value of incorporating decision-making feedback into forecasting.**

***Index Terms*—Prediction interval, robust optimization, deep learning, decision-informed, renewable energy, uncertainty**

I. Introduction

Renewable energy generation (REG) has experienced substantial growth in recent decades as an effective approach to reducing carbon emissions and mitigating global warming. By 2024, the global installed capacity of wind and photovoltaic (PV) power reached 2998 GW, with 564 GW new installation [1]. However, due to high dependency on meteorological conditions, REG exhibits inherent volatility and intermittency and introduces significant uncertainty to power systems. Accurate probabilistic forecasting, which provides valuable information about uncertainty of REG, becomes increasingly essential for power systems with large-scale integration of REG [2].

Early studies on probabilistic forecasting rely on predefined assumptions about the distribution of uncertainties, referred to as parametric probabilistic forecasting [3]. To facilitate high-quality probabilistic forecasting, nonparametric prediction intervals (PIs) of PV power generation are formulated as a pair of quantiles and optimized through quantile regression [4]. Furthermore, the quantile levels are set as decision variables to seek for the optimal PIs based on extreme learning machine (ELM) [5]. In [6], Pareto optimality between reliability and interval width of PIs for wind power forecasting is established to avoid specifying quantile levels.

This work was supported by the National Natural Science Foundation of China under Grant 52422706 and Grant 52277130.

Generally, the uncertainty information provided by probabilistic forecasting is relayed to downstream decision-making in power systems, such as stochastic optimization [7] and robust optimization [8]. However, forecasts with higher accuracy do not necessarily contribute to better decision-making due to the mismatch between the objectives of forecasting and decision. Beyond the statistical performance, growing research interest has been aroused toward the value of forecasting [9], [10]. Instead of minimizing the prediction error, an iterative learning approach for value-oriented point forecasting of REG is developed to reduce operational costs of the virtual power plant [11]. For decision-making under uncertainty, a task-based load forecasting model that incorporates the process of generation scheduling is established to minimize the cost of stochastic economic dispatch [7]. Another data-driven method based on the augmented Lagrangian method is proposed to learn uncertainty sets for robust reserve allocation [8]. However, the assumed Gaussian distribution of uncertainty in [7] and [8] hinders their further application in power systems with high-penetration REG. To address above issues, an integrated probabilistic forecasting and robust optimization method is elaborated based on machine learning by optimizing parameters of the probabilistic forecasting model and decision variables of robust optimization simultaneously [12].

To align the uncertainty quantification with the optimization objective of downstream decision-making, a novel bilevel optimization-based framework for decision-informed prediction intervals of renewable energy generation is firstly developed. In the proposed framework, interval forecasting is formulated as the upper-level problem with model parameters serving as decision variables, while downstream robust optimization is introduced as the lower-level problem. Instead of solely pursuing the best statistical performance, the interval forecaster explicitly minimizes the system operational cost with PIs coverage constraint. Furthermore, the robust optimal power flow (ROPF) is employed as the downstream task and reformulated as a linear program embedded within a convex optimization layer. An improved difference of convex functions-based surrogate is devised to approximate the intractable chance constraint, combined with an iterative solution strategy based on augmented Lagrangian method inspired by [8], which facilitates the end-to-end training of advanced deep learning models. Comprehensive numerical studies on the IEEE 14-bus and 39-bus systems with integrated wind power generation validate the superior performance of the proposed

979-8-3315-8138-1

method.

Main contributions of the paper are concluded as follows,

1) Decision-informed prediction intervals are firstly proposed to align the uncertainty quantification with the objective of downstream decision-making by formulating a bilevel optimization-based framework.
2) Robust optimal power flow is employed as the downstream task and reformulated as a differentiable linear program embedded within a convex optimization layer.
3) An improved difference of convex functions-based approximation is developed to substitute chance constraint and combined with augmented Lagrangian method to enable end-to-end training of deep learning models.

The remainder of the paper is organized as follows. Section II develops the bilevel optimization-based framework of decision-informed PIs. The linearization of ROPF and training algorithms of deep learning forecasting models are illustrated in Section III. Numerical studies are performed in Section IV. Finally, Section V concludes the work.

II. Framework of Decision-Informed Prediction Intervals

A. Preliminary of Prediction Intervals

Denoting the future renewable energy generation at time t as y_t, the forecaster aims to provide reliable prediction interval $\mathcal{I}_t^\beta$ given a nominal coverage proportion $100\,(1-\beta)\,\%$,

$$\mathbb{P}\big(y_t \in \mathcal{I}_t^\beta\big) = 100\,(1-\beta)\,\% \tag{1}$$

where $\mathcal{I}_t^\beta$ is defined by the lower and upper bounds ℓ_t and u_t.

Generally, calibration and sharpness are two fundamental criteria for evaluating PIs. Calibration indicates the reliability of PIs, which can be reflected by the deviation between the empirical coverage proportion (ECP) and nominal coverage proportion (NCP), also known as the average coverage deviation (ACD),

$$\mathrm{ACD} = \mathrm{ECP} - 100\,(1-\beta)\,\% \tag{2}$$

where ECP denotes the average coverage of prediction intervals, defined by,

$$\mathrm{ECP} = \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} \mathbb{I}\big(y_t \in \mathcal{I}_t^\beta\big) \tag{3}$$

PIs with the ACD value closer to 0 present higher reliability.

Sharpness indicates the concentration of PIs, commonly measured by the average width (AW),

$$\mathrm{AW} = \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} (u_t - \ell_t) \tag{4}$$

where $\mathcal{T}$ is the set of time indices, and $|\mathcal{T}|$ denotes the number of elements in $\mathcal{T}$. Calibrated PIs with lower AW value carry more valuable information and are preferred by forecasters.

Furthermore, the interval score (IS) [13] is introduced to assess the overall skill of PIs, defined as,

$$\begin{aligned}\mathrm{IS} = \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} \Big[&(u_t - \ell_t) + \frac{2}{\beta}(\ell_t - y_t)\,\mathbb{I}(y_t \le \ell_t) \\ &+ \frac{2}{\beta}(y_t - u_t)\,\mathbb{I}(u_t \le y_t)\Big]\end{aligned} \tag{5}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. It should be noted that the negatively oriented score arises from the evaluation of a pair of quantiles at central proportions, which tends to show bias toward non-central PIs.

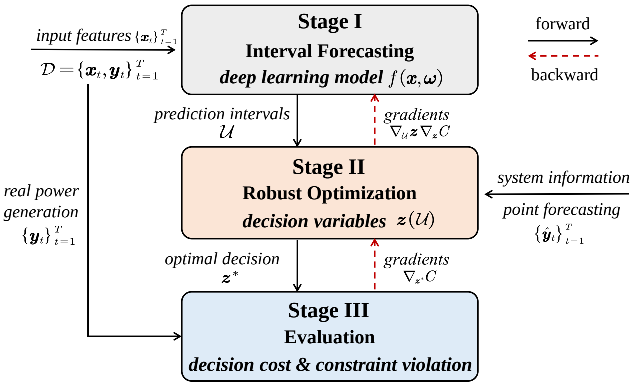


Fig. 1. Framework of decision-informed prediction intervals.

B. Decision-making with Uncertainty in Power Systems

To address uncertainty of future renewable energy generation, day-ahead robust optimization [12] is introduced as,

$$\min_{\boldsymbol{z}} C(\boldsymbol{z}) \tag{6a}$$

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{z},\boldsymbol{\xi}) \le \boldsymbol{0}, \forall \boldsymbol{\xi} \in \mathcal{U} \tag{6b}$$

$$\boldsymbol{h}(\boldsymbol{z},\boldsymbol{\xi}) = \boldsymbol{0} \tag{6c}$$

where $\boldsymbol{z}$ denotes the decision variables vector, C denotes the objective function, $\boldsymbol{\xi}$ is the random vector, $\mathcal{U}$ is the uncertainty set equivalent with the interval $\mathcal{I}_t^\beta$, $\boldsymbol{g}(\boldsymbol{z},\boldsymbol{\xi}) \le 0$ is the inequality constraints involving the uncertainty, and $\boldsymbol{h}(\boldsymbol{z}) = 0$ is the equality constraints. Given the uncertainty set $\mathcal{U}$ decided by bounds $\boldsymbol{u}_t$ and $\boldsymbol{\ell}_t$, the optimal decision plan $\boldsymbol{z}^*$ is obtained by solving (6), which is influenced by $\boldsymbol{u}_t$ and $\boldsymbol{\ell}_t$.

C. Bilevel Optimization-based Framework

In the paper, a bilevel optimization-based framework of decision-informed PIs, is proposed to enable the downstream tasks under uncertainty to guide the formulation of PIs, rather than focusing solely on statistical performance, as shown in Fig. 1. The bilevel optimization is expressed as,

$$\min_{\boldsymbol{\omega}} C(\boldsymbol{z}^*) \tag{7a}$$

$$\text{s.t. } (\boldsymbol{\ell},\boldsymbol{u}) = f(\boldsymbol{x},\boldsymbol{\omega}) \tag{7b}$$

$$\mathbb{P}(\boldsymbol{\ell} \le \boldsymbol{y} \le \boldsymbol{u}) \ge 100\,(1-\beta)\,\% \tag{7c}$$

$$\boldsymbol{z}^* = \arg\min C(\boldsymbol{z}) \text{ s.t. (6b)} - \text{(6c)} \tag{7d}$$

where $\boldsymbol{\omega}$ denotes the parameters of the deep learning model $f(\boldsymbol{x},\boldsymbol{\omega})$ which takes the features $\boldsymbol{x}$ as input and generates the lower and upper bounds of the uncertainty set $\mathcal{U}$.

In the proposed framework, the upper-level problem (7a) seeks to optimize the overall operational cost C, with model parameters $\boldsymbol{\omega}$ serving as decision variables. The uncertainty set $\mathcal{U}$, formulated through PIs generated by $f(\boldsymbol{x},\boldsymbol{\omega})$, is then conveyed to the lower-level problem. Upon solving the robust optimization (7d) and real renewable generation $\boldsymbol{y}$ is revealed, the model parameters $\boldsymbol{\omega}$ are updated using gradient descent.

III. Model Formulation and Training Algorithm

A. Robust Optimal Power Flow Model

The robust DCOPF, as one of the most common downstream tasks, is mainly investigated in the paper. The affine balancing

979-8-3315-8138-1

control strategy is utilized to handle the imbalance induced by the uncertainty of renewable energy generation. The concrete formulation of robust OPF at time index t is expressed as,

$$\min_{\mathcal{Z}_t} \boldsymbol{c}_{\mathrm{g}}^\top \hat{\boldsymbol{p}}_t^{\mathrm{g}} + \boldsymbol{c}_{\mathrm{u}}^\top \boldsymbol{r}_t^{\mathrm{u}} + \boldsymbol{c}_{\mathrm{d}}^\top \boldsymbol{r}_t^{\mathrm{d}} \tag{8a}$$

$$\text{s.t. } \mathbf{1}^\top \boldsymbol{p}_t^{\mathrm{g}} + \mathbf{1}^\top \boldsymbol{y}_t = \mathbf{1}^\top \boldsymbol{d}_t \tag{8b}$$

$$\hat{\boldsymbol{p}}_t^{\mathrm{g}} + \boldsymbol{r}_t^{\mathrm{u}} \leq \overline{\boldsymbol{p}}^{\mathrm{g}},\ \hat{\boldsymbol{p}}_t^{\mathrm{g}} - \boldsymbol{r}_t^{\mathrm{d}} \geq \underline{\boldsymbol{p}}^{\mathrm{g}} \tag{8c}$$

$$\boldsymbol{p}_t^{\mathrm{g}} = \hat{\boldsymbol{p}}^{\mathrm{g}} - \mathbf{1}^\top (\boldsymbol{y}_t - \hat{\boldsymbol{y}}_t) \boldsymbol{k} \tag{8d}$$

$$\mathbf{1}^\top \boldsymbol{k} = 1,\ \boldsymbol{k} \geq 0 \tag{8e}$$

$$-\boldsymbol{r}_t^{\mathrm{d}} \leq \boldsymbol{p}_t^{\mathrm{g}} - \hat{\boldsymbol{p}}_t^{\mathrm{g}} \leq \boldsymbol{r}_t^{\mathrm{u}},\ \forall \boldsymbol{y}_t \in \mathcal{U}_t \tag{8f}$$

$$-\overline{\boldsymbol{p}}^{\mathrm{l}} \leq \boldsymbol{T}\big[\boldsymbol{C}_g \boldsymbol{p}_t^{\mathrm{g}} + \boldsymbol{C}_y \boldsymbol{y}_t - \boldsymbol{C}_d \boldsymbol{d}_t\big] \leq \overline{\boldsymbol{p}}^{\mathrm{l}},\ \forall \boldsymbol{y}_t \in \mathcal{U}_t \tag{8g}$$

where $\boldsymbol{c}_{\mathrm{g}}$, $\boldsymbol{c}_{\mathrm{u}}$ and $\boldsymbol{c}_{\mathrm{d}}$ are the cost coefficient of scheduled conventional unit generation $\hat{\boldsymbol{p}}_t^{\mathrm{g}}$, the upward reserve provision $\boldsymbol{r}_t^{\mathrm{u}}$ and the downward reserve provision $\boldsymbol{r}_t^{\mathrm{d}}$. The scheduled unit generation $\hat{\boldsymbol{p}}_t^{\mathrm{g}}$ is determined according to the expected renewable energy generation $\hat{\boldsymbol{y}}_t$.

The objective (8a) aims to minimize the operation cost with $\mathcal{Z}_t = \left\{\hat{\boldsymbol{p}}_t^{\mathrm{g}}, \boldsymbol{r}_t^{\mathrm{u}}, \boldsymbol{r}_t^{\mathrm{d}}\right\}$ as the set of decision variables. The power balance constraint (8b) requests the power supply to be equal with the demand $\boldsymbol{d}_t$. The constraints (8c) and (8f) ensure the conventional unit output and adjustment would not exceed the generation interval $[\underline{\boldsymbol{p}}^g, \overline{\boldsymbol{p}}^g]$, the downward reserve limit $\boldsymbol{r}_t^{\mathrm{d}}$ and upward reserve limit $\boldsymbol{r}_t^{\mathrm{u}}$, respectively. The affine policy adopted to handle the uncertainty is realized by (8e), where $\boldsymbol{k}$ is the non-negative coefficient vector. In the paper, $\boldsymbol{k}$ is set according to the generator capacity for simplicity. The line flow constraints (8g) keep the power flow on transmission lines within the thermal limits $\overline{\boldsymbol{p}}^{\mathrm{l}}$, where $\boldsymbol{T}$ denotes the matrices of power transfer distribution factor (PTDF), and $\boldsymbol{C}_{\mathrm{g}}$, $\boldsymbol{C}_{\mathrm{y}}$, $\boldsymbol{C}_{\mathrm{d}}$ are the incidence matrices corresponding to the conventional generators, renewable energy generation and load nodes, respectively.

The robust constraints (8f)-(8g), which include finite variables and infinite number of constraints, can be reformulated by considering the worst scenarios [14]. Combining the affine policy (8d), the robust constraints (8f)-(8g) are equivalent to the following linear constraints,

$$-\mathbf{1}^\top (\boldsymbol{\ell}_t - \hat{\boldsymbol{y}}_t) \boldsymbol{k} \leq \boldsymbol{r}_t^{\mathrm{u}} \tag{9a}$$

$$\mathbf{1}^\top (\boldsymbol{u}_t - \hat{\boldsymbol{y}}_t) \boldsymbol{k} \leq \boldsymbol{r}_t^{\mathrm{d}} \tag{9b}$$

$$\boldsymbol{T}\big[\boldsymbol{C}_{\mathrm{g}} \left(\hat{\boldsymbol{p}}^{\mathrm{g}} + \mathbf{1}^\top \hat{\boldsymbol{y}}_t \boldsymbol{k}\right) - \boldsymbol{C}_{\mathrm{d}} \boldsymbol{d}_t\big] + \boldsymbol{M}^+ \boldsymbol{u}_t - \boldsymbol{M}^- \boldsymbol{\ell}_t \leq \overline{\boldsymbol{p}}^{\mathrm{l}} \tag{9c}$$

$$\boldsymbol{T}\big[\boldsymbol{C}_{\mathrm{g}} \left(\hat{\boldsymbol{p}}^{\mathrm{g}} + \mathbf{1}^\top \hat{\boldsymbol{y}}_t \boldsymbol{k}\right) - \boldsymbol{C}_{\mathrm{d}} \boldsymbol{d}_t\big] + \boldsymbol{M}^+ \boldsymbol{\ell}_t - \boldsymbol{M}^- \boldsymbol{u}_t \geq -\overline{\boldsymbol{p}}^{\mathrm{l}} \tag{9d}$$

$$\boldsymbol{M} = \boldsymbol{T}\boldsymbol{C}_{\mathrm{y}} - \boldsymbol{T}\boldsymbol{C}_{\mathrm{g}} \boldsymbol{k} \mathbf{1}^\top \tag{9e}$$

where $(\cdot)^+$ and $(\cdot)^-$ take the positive and negative parts of the corresponding elements, respectively. The $\boldsymbol{\ell}_t$ and $\boldsymbol{u}_t$ are generated by the interval forecaster $f(\boldsymbol{x}_t, \boldsymbol{\omega})$ taking features $\boldsymbol{x}_t$ as input. Hence, the investigated robust OPF is reformulated as a linear program, which can be written in a compact form as follows, which can be further integrated into deep learning models with the cvxpylayer [15].

$$\min_{\mathcal{Z}_t} \boldsymbol{c}^\top \boldsymbol{z}_t \tag{10a}$$

$$\text{s.t. } \boldsymbol{A}\boldsymbol{z}_t + \boldsymbol{b}_t \leq \mathbf{0} \tag{10b}$$

$$\boldsymbol{C}\boldsymbol{z}_t + \boldsymbol{d}_t = \mathbf{0} \tag{10c}$$

B. Decision-Informed Prediction Intervals for Robust OPF

Based on the proposed framework of decision-informed prediction intervals, the decision-informed PIs for robust OPF can be formulated as,

$$\min_{\boldsymbol{\omega}} \sum_{t \in \mathcal{T}} \boldsymbol{c}^\top \boldsymbol{z}_t^* \tag{11a}$$

$$\text{s.t. } (\boldsymbol{\ell}_t, \boldsymbol{u}_t) = f(\boldsymbol{x}_t, \boldsymbol{\omega}),\ \forall t \in \mathcal{T} \tag{11b}$$

$$\mathbb{P}(\boldsymbol{\ell} \leq \boldsymbol{y} \leq \boldsymbol{u}) \geq 100(1-\beta)\,\% \tag{11c}$$

$$\begin{aligned} \boldsymbol{z}_t^* = \arg\min \boldsymbol{c}^\top \boldsymbol{z}_t \\ \text{s.t. (10b)} - \text{(10c)} \end{aligned},\ \forall t \in \mathcal{T} \tag{11d}$$

The coverage constraint (11c) can be reformulated using the sample average approximation, expressed as,

$$\frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \mathbb{I}(\boldsymbol{\ell}_t \leq \boldsymbol{y}_t \leq \boldsymbol{u}_t) \geq 100(1-\beta)\,\% \tag{12}$$

By introducing an auxiliary variable η_t, the logical condition $\boldsymbol{\ell}_t \leq \boldsymbol{y}_t \leq \boldsymbol{u}_t$ involved in (12) is equivalent to $\eta_t \leq 0$,

$$\eta_t = \max\left[\max(\boldsymbol{y}_t - \boldsymbol{u}_t), \max(\boldsymbol{\ell}_t - \boldsymbol{y}_t)\right] \tag{13}$$

Since the indicator function $\mathbb{I}(\cdot)$ in the coverage constraint (12) results in discontinuity of the feasible region, an improved difference of two convex functions is adopted as a surrogate for the indicator function, formulated as,

$$\hat{\mathbb{I}}(\eta_t) = (\kappa\eta_t + 0.5)^+ - (\kappa\eta_t - 0.5)^+ \tag{14}$$

where the positive coefficient κ controls the similarity between the indicator function and the surrogate. Therefore, the coverage constraint (12) is reformulated as,

$$\frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \left[(\kappa\eta_t + 0.5)^+ - (\kappa\eta_t - 0.5)^+\right] \geq 1 - \beta \tag{15}$$

C. Iterative Solution Strategy for End-to-End Learning

In the paper, an iterative solution strategy is proposed to incorporate advanced deep learning models for addressing more complex tasks. Firstly, the approximated coverage constraint (15) is integrated into the objective function (11a) using the augmented Lagrangian method, shown as,

$$\mathcal{L} = \sum_{t \in \mathcal{T}} \boldsymbol{c}^\top \boldsymbol{z}_t^* + \lambda\big[h(\boldsymbol{y}, \boldsymbol{\ell}, \boldsymbol{u}) + s\big] + \frac{\sigma}{2}\big[h(\boldsymbol{y}, \boldsymbol{\ell}, \boldsymbol{u}) + s\big]^2 \tag{16}$$

$$h(\boldsymbol{y}, \boldsymbol{\ell}, \boldsymbol{u}) = (1-\beta) - \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \Big[(\kappa\eta_t + 0.5)^+ - (\kappa\eta_t - 0.5)^+\Big] \tag{17}$$

where λ denotes the Lagrange multiplier, σ represents the penalty coefficient, and s is a non-negative auxiliary variable. During the iterative training process, the optimal value of s can be deduced as follows [16] when parameters of the forecasting model $f(\boldsymbol{x}, \boldsymbol{\omega})$ are updated,

$$s = \max\left(0, -\frac{\lambda}{\sigma} - h(\boldsymbol{y}, \boldsymbol{\ell}, \boldsymbol{u})\right) \tag{18}$$

It can be deduced that the Lagrangian function $\mathcal{L}$ is convex with respect to $\boldsymbol{z}_t$, $\boldsymbol{\ell}_t$ and $\boldsymbol{u}_t$. The gradient of $\mathcal{L}$ with respect to the model parameters $\boldsymbol{\omega}$ is given by,

$$\nabla_{\boldsymbol{\omega}} \mathcal{L} = \nabla_{\boldsymbol{\omega}} \boldsymbol{z}_t^* \cdot \nabla_{\boldsymbol{z}_t^*} \mathcal{L} + \nabla_{\boldsymbol{\omega}} \boldsymbol{\ell}_t \cdot \nabla_{\boldsymbol{\ell}_t} \mathcal{L} + \nabla_{\boldsymbol{\omega}} \boldsymbol{u}_t \cdot \nabla_{\boldsymbol{u}_t} \mathcal{L} \tag{19}$$

Since the lower-level problem (11d) is a linear program, the cvxpylayer [15] is employed to solve (11d) and obtain both the optimal value $\boldsymbol{z}_t^*$ and its gradient $\nabla_{\boldsymbol{\omega}} \boldsymbol{z}_t^*$. The detailed training procedure of the deep learning model $f(\boldsymbol{x}, \boldsymbol{\omega})$ is presented in Algorithm 1.

Algorithm 1 Training Algorithm for Decision-Informed PIs

Input: training dataset $\mathcal{D}_{\text{train}} = \{\boldsymbol{x}_t, y_t\}_{t \in \mathcal{T}_{train}}$; learning rate ρ; initial parameters λ_0, $\boldsymbol{\omega}_{0,0}$;
Output: optimal model parameters $\boldsymbol{\omega}$
for epoch $i = 1, \ldots, i^{\max}$ **do**
  for batch $j = 1, \ldots, j^{\max}$ **do**
    $\mathcal{U}_j \leftarrow f\left(\boldsymbol{x}_j, \boldsymbol{\omega}_{i-1,j-1}\right)$
    Solve (11d) and calculate $\nabla_{\boldsymbol{\omega}} \boldsymbol{z}_t^*$
    Calculate $\nabla_{\boldsymbol{\omega}} \mathcal{L}$ according to (19)
    $\boldsymbol{\omega}_{i-1,j} \leftarrow \boldsymbol{\omega}_{i-1,j-1} - \rho \nabla_{\boldsymbol{\omega}} \mathcal{L}$
  $\boldsymbol{\omega}_{i,0} \leftarrow \boldsymbol{\omega}_{i-1,j^{\max}}$
  $\lambda_i \leftarrow \lambda_{i-1} + \sigma \frac{1}{j^{\max}} \sum_{j=1}^{j^{\max}} h\left(\boldsymbol{y}_j, \boldsymbol{\ell}_j, \boldsymbol{u}_j\right)$
return $\boldsymbol{\omega}_{i^{\max},0}$

TABLE I
PERFORMANCE OF DIFFERENT FORECASTING METHODS

Methods	ACD	ECP	AW	IS	Cost ($)
BELM	3.66%	93.66%	0.6228	0.7144	7682.27
QRPI	0.32%	90.32%	0.5873	0.6570	7444.19
QDPI	3.44%	93.44%	0.5447	0.6690	7415.53
MOLUBE	1.29%	91.29%	0.5734	0.6600	7376.41
DCUS	2.12%	92.12%	0.5196	0.6650	7082.77
DIPI	0.62%	90.62%	0.5510	0.7500	6910.68

IV. NUMERICAL STUDIES

A. Description of Experiments

Numerical studies on the IEEE 14-bus system from [17] are conducted to validate the performance of the proposed method. The forecasting task is day-ahead wind power forecasting, while the decision task is day-ahead robust optimal power flow. A wind farm with 150 MW capacity is integrated at bus 5. The normalized wind power generation data are derived from the first wind farm in the Global Energy Forecasting Competition 2014 (GEFCom2014) [2]. The time ranges from January 2012 to August 2013 with 1-hour resolution. The dataset is divided into two subsets in a 2:1 ratio for developing point and interval forecasting models with 40% of the second subset reserved for evaluating interval forecasting performance.

The proposed decision-informed PIs (DIPI) method is compared with four predict-then-optimize methods, including the bootstrap-based extreme learning machine (BELM) [3], quantile regression-based PIs (QRPI) [4], [18], quality-driven PIs (QDPI) [19] and multi-objective lower and upper bounds estimation (MOLUBE) [6], as well as a predict-and-optimze method, namely the decision-centric uncertainty set (DCUS) [20] The deep neural network with Gaussian noise layers [18], is adopted as the base model for deep learning-based methods.

B. Comparison on Statistical Score and Operational Cost

The statistical forecasting performance and the resulting operational costs of different interval forecasting approaches under 90% NCP together are summarized in Table I. Among the benchmarks, the QRPI method presents the smallest ACD and IS values. However, the better statistical performance indicated by IS does not necessarily result in improved operational decisions. In contrast, the two decision-aware approaches DCUS and DIPI yield more economic operational plans than other four predict-then-optimize methods. Specifically, the proposed DIPI exhibits a balanced statistical performance with an ACD of 0.62% and a reasonably compact interval width, while achieving a substantial cost reduction of at least 6.31% compared with the predict-and-optimize benchmarks and 2.43% compared with DCUS. This highlights that DIPI effectively aligns interval construction with downstream decision objectives rather than solely pursuing statistical optimality. As illustrated in Fig. 2, DIPI achieves tighter prediction intervals to reflect rapid variations of wind power, facilitating the robust OPF to allocate more economic reserves.

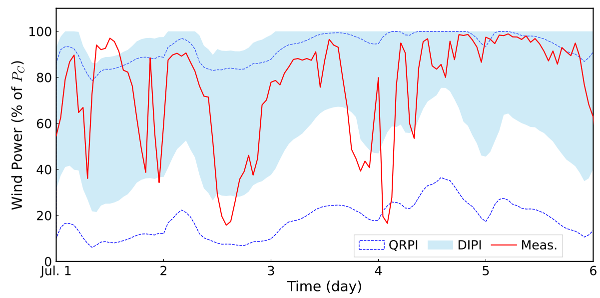
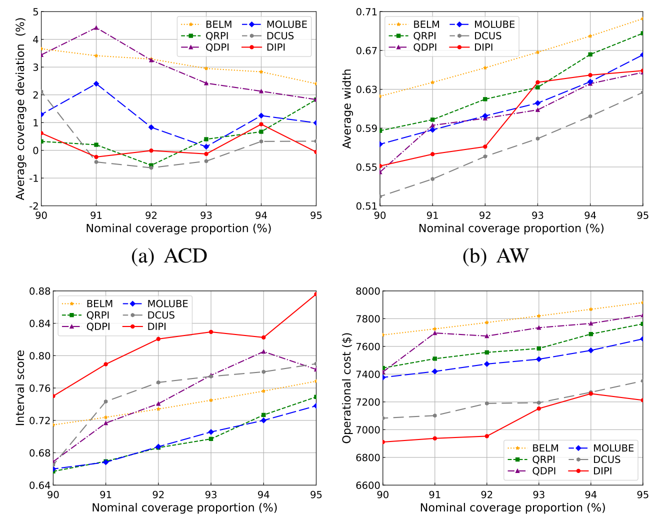


Fig. 2. Wind power forecasting curves generated by TDPPF and benchmarks.



(a) ACD (b) AW

(c) IS (d) Cost

Fig. 3. Performance comparison under different NCPs.

C. Analysis on Different Nominal Coverage Proportions

Sensitivity analysis of interval forecasting methods across different NCPs ranging from 90% to 95% is conducted and displayed in Fig. 3. It is observed that both the width of PIs and operational cost increases as the NCP rises. Regarding reliability indicated by ACD, the ACD curve of the proposed DIPI method is the closest to zero with the absolute ACD value limited within 1%, validating the improved difference of convex functions-based approximation used in DIPI. Moreover, the operational cost associated with DIPI consistently remains the lowest across all NCPs, demonstrating the robustness of the proposed method. The minimal cost

979-8-3315-8138-1

TABLE II
PERFORMANCE OF DIFFERENT FORECASTING METHODS ON THE IEEE 118-BUS SYSTEM

Methods	ACD	ECP	Volume	Cost ($)
QRPI	-6.14%	83.86%	0.6258	122347.42
QDPI	3.13%	93.13%	0.3331	118888.20
DIPI	-1.14%	88.86%	0.3741	115930.28

reduction of DIPI compared to predict-then-optimze benchmarks ranges from 4.12% to 6.96%. Although the QRPI and MOLUBE method achieves the best interval score, their corresponding operational costs are still significantly higher than those of the decision-informed PIs, further emphasizing the value of feedback from decision-makers to forecasters for better decisions. The decision-aware DCUS method achieves the shortest interval width but incurs higher operational cost than the proposed DIPI method since it only optimizes the unconditional uncertainty set of prediction errors.

D. Experiments on the Large System

The IEEE 118-bus system is selected to validate the scalability of the DIPI method. Five wind farms with 300 MW capacity are integrated at buses 10, 32, 66, 89 and 101, respectively. The normalized wind power generation data originates from the first five wind farms in the GEFCom2014 dataset [2]. The NCP is set at 90%. In the multivariate case, the interval width in the loss function of QDPI is replaced by the volume of prediction boxes. The QRPI method is extended to generate prediction boxes by assuming the wind power generation of different farms to be independent. Performance of different methods is detailed in Table II. It is evident that the neglect of dependencies in QRPI leads to poor coverage and the highest operational costs. The prediction boxes generated by DIPI reduce the operational costs of QRPI and QDPI by 5.24% and 2.49%, respectively, emphasizing the value of DIPI in large systems. The computational efficiency of DIPI is tested on a computer with Intel Xeon Platinum 8481C CPU@3.8GHz. Training time per epoch is 80.78s for the IEEE 14-bus system and 376.27s for the IEEE 118-bus system. Once the training is completed, the inference time is on the order of milliseconds, which is adequate for practical applications.

V. CONCLUSION

This paper presents a novel decision-informed framework for constructing prediction intervals that directly incorporate the objective of downstream robust optimal power flow. By formulating the forecasting task and downstream decision-making as a unified bilevel optimization, the proposed method aligns the structure of prediction intervals with the operational target of ROPF rather than solely relying on statistical performance. The downstream task ROPF is reformulated as a differentiable linear program and embedded into deep learning models as a convex optimization layer. The improved surrogate for the PIs coverage constraint combined with the augmented Lagrangian method, enables gradient-based end-to-end training of deep learning models. Experiments on the IEEE 14-bus and 118-bus systems demonstrate that the proposed DIPI generates more reliable prediction intervals which yields significantly lower operational costs under various coverage proportions. The results underscore the importance of coupling forecasting with decision-making to improve operational performance.

REFERENCES

[1] IRENA(2025), "Renewable capacity statistics 2025," International Renewable Energy Agency, Abu Dhabi.

[2] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond," *Int. J. Forecasting*, vol. 32, no. 3, pp. 896–913, Jul. 2016.

[3] 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.

[4] C. Wan, J. Lin, Y. Song, Z. Xu, and G. Yang, "Probabilistic forecasting of photovoltaic generation: An efficient statistical approach," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2471–2472, May 2017.

[5] C. Zhang and Y. Fu, "Probabilistic electricity price forecast with optimal prediction interval," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 442–452, Jan. 2024.

[6] Y. Chen, S. S. Yu, C. P. Lim, and P. Shi, "Multi-objective estimation of optimal prediction intervals for wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 15, no. 2, pp. 974–985, Apr. 2024.

[7] J. Han, L. Yan, and Z. Li, "A task-based day-ahead load forecasting model for stochastic economic dispatch," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5294–5304, Nov. 2021.

[8] R. Mieth and H. V. Poor, "Prescribed robustness in optimal power flow," *Electr. Power Syst. Res.*, vol. 235, p. 110704, Oct. 2024.

[9] H. Zhang, R. Li, Q. Du, J. Tao, S. Pineda, G. Kariniotakis, S. Camal, C. B. Monroc, M. Sun, C. Wan, W. Xu, and F. Teng, "Decision-focused learning for power system decision-making under uncertainty," *IEEE Trans. Power Syst.*, pp. 1–18, 2025.

[10] M. Beykirch, A. Bott, T. Janke, and F. Steinke, "The value of probabilistic forecasts for electricity market bidding and scheduling under uncertainty," *IEEE Trans. Power Syst.*, vol. 39, no. 6, pp. 6986–6997, Nov. 2024.

[11] Y. Zhang, M. Jia, H. Wen, Y. Bian, and Y. Shi, "Toward value-oriented renewable energy forecasting: An iterative learning approach," *IEEE Trans. Smart Grid*, pp. 1–1, 2024.

[12] C. Wan, K. Zhang, C. Zhao, H. Liu, and P. Ju, "An integrated probabilistic forecasting and robust optimization method for ac-opf under high-penetration renewable energy," *IEEE Trans. Sustain. Energy*, pp. 1–15, 2025.

[13] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *J. Am. Stat. Assoc.*, vol. 102, no. 477, pp. 359–378, Mar. 2007.

[14] Y. Wang, Z. Zhou, A. Botterud, and K. Zhang, "Optimal wind power uncertainty intervals for electricity market operation," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 199–210, Jan. 2018.

[15] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and J. Z. Kolter, "Differentiable convex optimization layers," in *Advances in Neural Information Processing Systems*, vol. 32. Curran Associates, Inc., 2019.

[16] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 2006.

[17] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang *et al.*, "The power grid library for benchmarking ac optimal power flow algorithms," Jan. 2021.

[18] W. Zhang, H. Quan, and D. Srinivasan, "An improved quantile regression neural network for probabilistic load forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4425–4434, Jul. 2019.

[19] T. Pearce, A. Brintrup, M. Zaki, and A. Neely, "High-quality prediction intervals for deep learning: A distribution-free, ensembled approach," in *Proceedings of the 35th International Conference on Machine Learning*. PMLR, Jul. 2018, pp. 4075–4084.

[20] J. Zhu, Y. Xu, J. Wang, J. Watson, J. Song, and H. Sun, "Decision-centric uncertainty sets for storage arbitrage: Integrating estimation and optimization," *IEEE Trans. Power Syst.*, vol. 40, no. 6, pp. 5511–5514, Nov. 2025.

979-8-3315-8138-1