

Article

Context-Aware Asymmetric Conformal Calibration of Renewable-Power Prediction Intervals for Day-Ahead Operational Risk Assessment

Peng Han ¹, Jun Zhao ², Yu Liu ¹, Xuehai Yu ³, Yizhou Wang ^{4,*} and Ran Li ⁴¹ Electric Power Dispatching and Control Center, State Grid Shanxi Electric Power Co., Ltd., Taiyuan 030001, China² Electric Power Research Institute, State Grid Shanxi Electric Power Co., Ltd., Taiyuan 030001, China³ Datong Power Supply Company, State Grid Shanxi Electric Power Co., Ltd., Datong 037008, China⁴ School of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

* Correspondence: wyz0901@sjtu.edu.cn

Abstract

High renewable penetration makes day-ahead operation sensitive to the directional effects of wind and photovoltaic forecast errors. Conventional prediction intervals mainly evaluate coverage and sharpness, but lower- and upper-boundary violations correspond to different operational risks: shortage-side supply-adequacy pressure and accommodation-side curtailment pressure. This paper proposes context-aware asymmetric conformal quantile regression (CA-ACQR) to construct directional renewable-power risk intervals. The method builds separate conformal scores for the two interval sides, estimates context-dependent boundary corrections, and reallocates the tail-risk budget under supply-priority, balanced, and accommodation-priority profiles. Case studies use regional wind and photovoltaic power data, with contextual groups defined by renewable type, lead-time block, forecast difficulty, weather-risk regime, and output level. CA-ACQR increases the prediction interval coverage probability (PICP) from 91.11% to 94.07% and reduces the accommodation-side violation rate from 4.37% to 1.44%. The results demonstrate selectable directional risk postures and quantify trade-offs among interval width, directional violations, normalized stress cost, and the 95% conditional value-at-risk stress cost.

Keywords: renewable power forecasting; prediction interval calibration; conformal quantile regression; asymmetric risk allocation; day-ahead operation; renewable accommodation



Academic Editor: Antonino Laudani

Received: 24 June 2026

Revised: 21 July 2026

Accepted: 24 July 2026

Published: 30 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the increasing penetration of wind and photovoltaic (PV) generation, renewable-power uncertainty has become a critical factor affecting day-ahead operation. Since renewable output is not yet realized at the scheduling stage, system operators must determine generation schedules, reserve margins, balancing resources, and market submissions in advance, making forecast errors directly relevant to both power balance and operating cost. More importantly, renewable forecast errors have clear directional operational implications. When actual output falls below the forecasted level, the system may face shortage-side pressure, increased reserve activation, and supply-adequacy risks, which have been closely related to unit commitment and reserve scheduling in wind-rich systems [1] and to spinning-reserve requirement estimation under wind uncertainty [2]. By contrast, when actual output exceeds the forecasted level, downward-regulation capability

may become insufficient, renewable accommodation may become difficult, and surplus-side pressure may appear in the form of wind and solar curtailment, reflecting the balancing challenges introduced by variable renewable generation [3] and the practical curtailment events observed in power systems with high renewable penetration [4]. Therefore, day-ahead operation requires not only a characterization of the uncertainty range of renewable output, but also an identification of the risk direction embedded in the prediction interval. In this sense, the lower boundary of the interval can be interpreted as a shortage-side supply-adequacy risk boundary, whereas the upper boundary can be interpreted as an accommodation-side renewable-accommodation risk boundary.

Renewable-power forecasting has gradually evolved from deterministic point prediction to probabilistic forecasting, in which future output uncertainty is commonly represented by quantiles, prediction intervals, probability densities, or scenario sets [5]. Existing studies have shown that uncertainty information plays an important role in the operational management of wind-power integration [6], while probabilistic forecasting theory provides the statistical foundation for calibrated predictive distributions and uncertainty-aware decision support [7]. A variety of methods have been developed for renewable-power probabilistic forecasting, including parametric error modeling, nonparametric prediction intervals, quantile regression, kernel-density estimation, numerical-weather-prediction ensemble post-processing, and machine learning models. Among them, quantile regression provides a classical statistical framework for conditional quantile estimation [8] and has been widely used in nonparametric probabilistic forecasting of wind power [9]. Optimal and nonparametric prediction intervals have also been developed for wind-power generation uncertainty [10], and related linear-programming formulations have been proposed for wind-power prediction interval construction [11]. For solar generation, Bayesian model averaging provides another route for probabilistic power forecasting [12]. Prediction intervals are typically evaluated in terms of reliability and sharpness, where the former reflects empirical coverage and the latter measures interval compactness. In recent years, conformal prediction has provided a distribution-free route for interval calibration [13], and distribution-free predictive inference has extended this idea to regression settings [14]. Conformalized quantile regression (CQR) further combines conditional quantile modeling with finite-sample coverage adjustment [15], while jackknife-based predictive inference provides a related refinement for uncertainty quantification [16].

Although existing probabilistic forecasting methods provide a solid basis for representing renewable-power uncertainty, the connection between prediction interval quality and day-ahead operational risk still requires further development. Conventional prediction intervals mainly focus on overall coverage and interval width, whereas in power-system operation, lower-boundary and upper-boundary violations correspond to different corrective actions, economic consequences, and operational risks. A lower-boundary violation indicates that actual renewable output is below the lower limit of the interval, requiring upward reserves, conventional-unit ramping, or other balancing resources to compensate for the power deficit; therefore, it essentially represents shortage-side supply-adequacy pressure. This interpretation is consistent with reserve scheduling and reserve adequacy analysis under wind uncertainty [1,2]. By contrast, an upper-boundary violation indicates that actual renewable output is above the upper limit of the interval, which may lead to insufficient downward-regulation resources, constrained renewable accommodation, and wind or solar curtailment risk; therefore, it essentially represents accommodation-side operational pressure. This surplus-side interpretation is closely related to balancing requirements under variable renewable generation [3] and to curtailment experience and prediction studies in renewable-rich systems [17]. Moreover, the system tolerance to these two types of directional risks varies with operating contexts such as renewable type, lead

time, prediction difficulty, and weather-risk regime. Hence, renewable-power prediction intervals for day-ahead operation should not be regarded merely as statistical coverage bands, but should be further developed into directional operating risk boundaries that can support tail-risk allocation under different operational preferences.

To this end, this paper proposes a context-aware asymmetric conformal quantile regression method, denoted as CA-ACQR, for constructing renewable-power directional risk intervals. The proposed method takes a base quantile forecasting model as input, treats lower- and upper-boundary violations as shortage-side and accommodation-side calibration signals, respectively, and estimates boundary corrections by incorporating context information such as renewable type, lead time, prediction difficulty, and weather-risk regime. On this basis, a profile-dependent tail-risk allocation mechanism is introduced, allowing the prediction interval to switch among supply-priority, balanced, and accommodation-priority operating modes. Case studies on a regional wind and PV dataset show that the proposed method yields interpretable trade-offs among directional violation rates, interval width, normalized stress cost, and the 95% conditional value at risk (CVaR₉₅) stress cost, thereby supporting day-ahead renewable-power risk assessment, reserve-aware analysis, and accommodation-risk screening.

2. Directional Risk Intervals for Renewable-Power Forecasting

2.1. Day-Ahead Renewable-Power Forecasting Setting

In day-ahead operation, renewable-power forecasting uses information available before real-time operation to characterize wind and PV uncertainty over future dispatch periods. The renewable output at target time t is normalized by the installed capacity or aggregation base so that interval metrics are comparable across station sizes and renewable types:

$$y_t = \frac{P_t}{P^{\max}}, \quad y_t \in [0, 1]. \quad (1)$$

Here, P_t is the observed renewable-power output at forecasting instance t , and $P^{\max}$ is the corresponding installed capacity or aggregation base. Equation (1) therefore maps the observed output to the common range $[0, 1]$.

The day-ahead information vector $\mathbf{x}_t$ includes historical renewable output, numerical weather prediction (NWP)-derived variables, time features, renewable type, and lead-time indicators. These inputs follow the operational information structure commonly used in wind-power forecasting [6]. For PV generation, probabilistic solar forecasting studies also emphasize weather-dependent inputs and forecast post-processing [12]. The base probabilistic forecaster produces uncalibrated quantiles and the associated initial interval as

$$\begin{aligned} \hat{q}_{\tau,t}^0 &= f_{\tau}(\mathbf{x}_t), \quad \tau \in (0, 1), \\ \hat{I}_t^0 &= [\hat{L}_t^0, \hat{U}_t^0] = [\hat{q}_{\alpha_L,t}^0, \hat{q}_{1-\alpha_U,t}^0], \\ \alpha &= \alpha_L + \alpha_U. \end{aligned} \quad (2)$$

Here, t indexes the forecasting samples, $\mathbf{x}_t$ is the day-ahead feature vector, and $f_{\tau}(\cdot)$ is the base forecasting model trained for quantile level τ . The output $\hat{q}_{\tau,t}^0$ is the corresponding uncalibrated conditional-quantile estimate, where the superscript 0 denotes the interval before conformal calibration. The initial lower and upper boundaries are $\hat{L}_t^0 = \hat{q}_{\alpha_L,t}^0$ and $\hat{U}_t^0 = \hat{q}_{1-\alpha_U,t}^0$, respectively. The quantities α_L and α_U are the lower- and upper-tail miscoverage probabilities, $\alpha = \alpha_L + \alpha_U$ is the total miscoverage rate, and $1 - \alpha$ is the nominal coverage probability. Equation (2) thus constructs the initial uncalibrated prediction interval from the lower- and upper-tail quantiles produced by the base probabilistic forecasting

model. A central interval is obtained when $\alpha_L = \alpha_U$; otherwise, the two sides receive different portions of the same total miscoverage budget.

2.2. Directional Interpretation of Prediction Interval Boundaries

A conventional prediction interval describes the statistical range in which renewable output may fall. In day-ahead operation, the two sides of this range imply different operating actions. Output below the lower boundary indicates unexpected renewable shortage and can increase upward reserve use, conventional-unit ramping pressure, and supply-adequacy risk. Output above the upper boundary indicates unexpected renewable surplus and can increase downward-regulation pressure, renewable accommodation difficulty, and curtailment exposure.

For the calibrated interval, the directional violation events and violation magnitudes are defined jointly as

$$\begin{aligned}\hat{I}_t &= [\hat{L}_t, \hat{U}_t], \\ v_t^L &= \mathbb{I}(y_t < \hat{L}_t), \quad v_t^U = \mathbb{I}(y_t > \hat{U}_t), \\ d_t^L &= [\hat{L}_t - y_t]_+, \quad d_t^U = [y_t - \hat{U}_t]_+.\end{aligned}\quad (3)$$

where $\hat{I}_t$ is the calibrated interval, $\hat{L}_t$ and $\hat{U}_t$ are its lower and upper boundaries, $\mathbb{I}(\cdot)$ is the indicator function, and $[x]_+ = \max(x, 0)$. The binary variables v_t^L and v_t^U identify lower- and upper-boundary violations, while d_t^L and d_t^U measure their normalized magnitudes. Equation (3) therefore separates whether a directional violation occurs from how far the observed output lies outside the corresponding boundary. The lower-boundary quantities describe shortage-side exposure, whereas the upper-boundary quantities describe accommodation-side exposure. Under supply stress or insufficient upward flexibility, stronger lower-side protection is preferred. Under renewable surplus, low load, or insufficient downward flexibility, stronger upper-side protection is preferred. Figure 1 illustrates this directional interpretation.

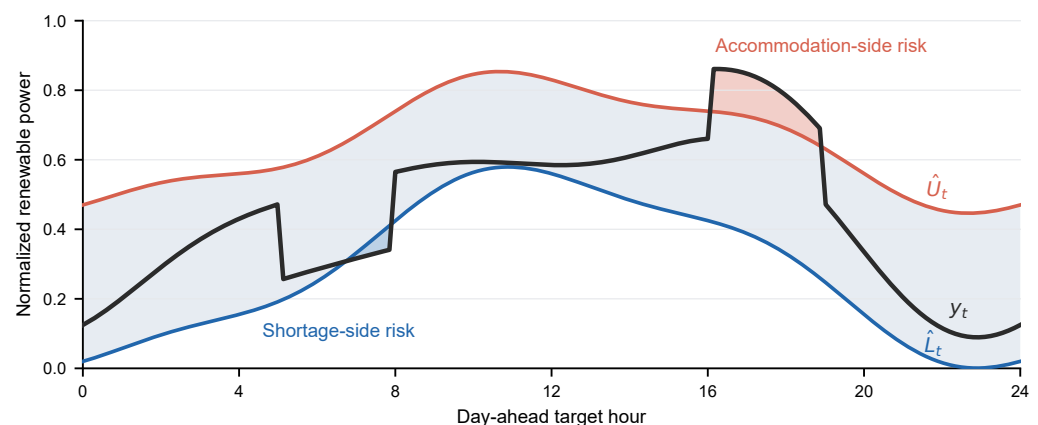


Figure 1. Schematic illustration of a calibrated renewable-power prediction interval. Here, y_t denotes the observed normalized renewable-power output at forecasting instance t . The lower boundary $\hat{L}_t$ is related to renewable shortfall and upward-reserve pressure, whereas the upper boundary $\hat{U}_t$ is related to renewable surplus and accommodation pressure. The trajectories illustrate the method and do not represent a specific test day; shaded areas mark the two directional violations.

2.3. Evaluation Metrics

Directional risk intervals are evaluated through reliability, sharpness, directional violation rates, and normalized operating pressure. The reliability and sharpness terminology follows the general theory of proper probabilistic forecasting evaluation [18]. Prediction-interval evaluation in wind-power forecasting uses the same empirical coverage and width

logic [10]. Reliability is measured by the prediction interval coverage probability (PICP), while sharpness is measured by the prediction interval normalized average width (PINAW):

$$\begin{aligned} \text{PICP} &= \frac{1}{N} \sum_{t=1}^N \mathbb{I}(\hat{L}_t \leq y_t \leq \hat{U}_t), \\ \text{PINAW} &= \frac{1}{NR_y} \sum_{t=1}^N (\hat{U}_t - \hat{L}_t). \end{aligned} \quad (4)$$

Here, N is the number of test samples, y_t is the observed normalized output, $\hat{L}_t$ and $\hat{U}_t$ are the calibrated boundaries, and $\mathbb{I}(\cdot)$ equals one when its condition is satisfied and zero otherwise. The normalization range is $R_y = y_{\max} - y_{\min}$; because renewable output is normalized to $[0, 1]$ in Equation (1), $R_y = 1$ in the case study. At a fixed nominal target $1 - \alpha$, higher PICP indicates stronger empirical coverage, whereas lower PINAW indicates sharper intervals. The two metrics are therefore interpreted jointly as a coverage–sharpness trade-off. Since overall coverage does not identify which boundary is violated, the directional violation rates are computed separately:

$$\begin{aligned} \text{VR}^L &= \frac{1}{N} \sum_{t=1}^N \mathbb{I}(y_t < \hat{L}_t), \\ \text{VR}^U &= \frac{1}{N} \sum_{t=1}^N \mathbb{I}(y_t > \hat{U}_t). \end{aligned} \quad (5)$$

The quantities VR^L and VR^U denote the shortage-side and accommodation-side violation rates, respectively. Equation (5) reveals how the total miscoverage is distributed between the two operational risk directions.

To compare directional exposure under consistent assumptions, a normalized stress-cost indicator is used. The shortage-side component is motivated by adequacy and reliability-cost studies, including value-of-lost-load reviews [19] and service-reliability estimates [20]. The accommodation-side component reflects balancing pressure from variable renewables [3] and the operational relevance of wind and solar curtailment [4]:

$$\begin{aligned} C_t^L &= c_L [\hat{L}_t - y_t]_+, \quad C_t^U = c_U [y_t - \hat{U}_t]_+, \\ C_t &= C_t^L + C_t^U, \quad \bar{C} = \frac{1}{N} \sum_{t=1}^N C_t. \end{aligned} \quad (6)$$

Here, c_L and c_U are normalized penalties for shortage-side and accommodation-side violation magnitudes. The terms C_t^L and C_t^U are the corresponding sample-level components, C_t is their sum, and $\bar{C}$ is the mean stress cost over the N test samples. In the case study, fixed symmetric normalized penalties are used, i.e., $c_L = c_U = 1$. The stress cost is therefore a diagnostic score based on directional violation magnitudes under common weights; it is not a physical dispatch cost and does not include interval width directly. Interval width is reported separately by PINAW. The upper tail of the stress-cost distribution is measured using value at risk (VaR) and CVaR:

$$\begin{aligned} \text{VaR}_{95}(C) &= \inf\{c : \Pr(C_t \leq c) \geq 0.95\}, \\ \text{CVaR}_{95}(C) &= \mathbb{E}[C_t \mid C_t \geq \text{VaR}_{95}(C)]. \end{aligned} \quad (7)$$

Equation (7) defines VaR_{95} as the 95th percentile of sample-level stress cost and CVaR_{95} as the mean within the largest 5% of costs. Thus, CVaR describes the most severe stress-cost outcomes rather than average performance. The same stress-cost weights and test samples

are used for all compared methods and operating profiles. Table 1 summarizes the metrics used in the case study.

Table 1. Evaluation metrics for directional renewable-power risk intervals.

Metric	Risk Meaning
PICP	Overall interval reliability
PINAW	Normalized interval sharpness
VR^L	Shortage-side violation rate
VR^U	Accommodation-side violation rate
Mean stress cost	Average operating pressure under fixed weights
CVaR ₉₅ stress cost	Tail operating pressure

3. Context-Aware Asymmetric Conformal Quantile Regression

3.1. Base Quantile Forecasting Model

CA-ACQR is designed as a post-processing and calibration framework rather than as a new base forecasting architecture, so that it can be integrated with existing probabilistic renewable-power forecasting models without changing their internal structures. The framework follows the modular use of quantile models, in which conditional quantiles are learned as forecasting primitives [8]. In renewable-power forecasting, direct quantile regression has provided an established implementation path for producing such probabilistic outputs [9]. Within the proposed framework, the base model first learns the conditional quantile structure from day-ahead information, while the conformal calibration layer subsequently adjusts the interval boundaries according to calibration samples, context conditions, and operating profiles. For a given feature vector, the base quantile and interval outputs are expressed as

$$\begin{aligned}\hat{q}_{\tau,t}^0 &= f_{\tau}(\mathbf{x}_t), \quad \tau \in (0,1), \\ \hat{I}_t^0 &= [\hat{L}_t^0, \hat{U}_t^0] = [\hat{q}_{\alpha_L,t}^0, \hat{q}_{1-\alpha_U,t}^0].\end{aligned}\quad (8)$$

Equation (8) restates the base-model interface used by the calibration framework: f_{τ} maps the day-ahead features $\mathbf{x}_t$ to the uncalibrated conditional quantile $\hat{q}_{\tau,t}^0$, and the two selected tail quantiles form $\hat{I}_t^0$. Because the proposed calibration layer only requires these quantile forecasts as inputs, it can be coupled with quantile regression, gradient boosting trees, random forests, neural networks, or any other learning model capable of producing conditional quantile estimates.

3.2. Asymmetric Conformal Scores

Although the base interval provides a preliminary description of the conditional distributional structure of renewable output, its empirical coverage may still deviate from the nominal level because of model error, finite-sample bias, and changes in forecast conditions. Conformal prediction addresses this calibration issue by constructing data-driven correction terms from held-out calibration scores [13], and distribution-free predictive inference provides the regression perspective for finite-sample uncertainty calibration [14]. CQR further connects conformal calibration with conditional quantile models, thereby enabling finite-sample coverage adjustment for quantile-based prediction intervals [15].

For the global CQR benchmark, the two interval sides are pooled into one nonnegative score and one shared correction:

$$\begin{aligned}s_t^G &= \max\{[\hat{q}_{\alpha_L,t}^0 - y_t]_+, [y_t - \hat{q}_{1-\alpha_U,t}^0]_+\}, \\ \Delta^G &= Q_{1-\alpha}(\{s_t^G : t \in \mathcal{C}\}), \\ \hat{I}_t^G &= [\hat{q}_{\alpha_L,t}^0 - \Delta^G, \hat{q}_{1-\alpha_U,t}^0 + \Delta^G].\end{aligned}\quad (9)$$

Here, $\mathcal{C} = \{(\mathbf{x}_t, y_t)\}_{t=1}^{n_c}$ is the held-out calibration set with n_c samples, s_t^G is the pooled global score, and $Q_\gamma(\cdot)$ denotes the empirical γ -quantile. The shared correction Δ^G expands both interval sides equally. In the reported experiment, $\Delta^G = 0$ because the uncalibrated interval already satisfied the nominal calibration requirement, which explains the identical Base-QR and CQR results.

CA-ACQR instead calibrates the lower and upper sides separately because their violations correspond to different operational consequences. The side-specific conformal scores on $\mathcal{C}$ are defined as

$$\begin{aligned} s_t^L &= [\hat{q}_{\alpha_L, t}^0 - y_t]_+, \\ s_t^U &= [y_t - \hat{q}_{1-\alpha_U, t}^0]_+. \end{aligned} \quad (10)$$

The nonnegative scores s_t^L and s_t^U quantify the shortage-side and accommodation-side exceedance magnitudes for calibration sample t , respectively. Based on these two score sets, the side-specific correction terms and the calibrated interval are obtained as

$$\begin{aligned} \Delta^L &= Q_{1-\alpha_L}(\{s_t^L : t \in \mathcal{C}\}), \quad \Delta^U = Q_{1-\alpha_U}(\{s_t^U : t \in \mathcal{C}\}), \\ \hat{L}_t &= \hat{q}_{\alpha_L, t}^0 - \Delta^L, \quad \hat{U}_t = \hat{q}_{1-\alpha_U, t}^0 + \Delta^U, \\ \hat{I}_t &= [\hat{L}_t, \hat{U}_t]. \end{aligned} \quad (11)$$

For side $k \in \{L, U\}$, the finite-sample quantile $Q_{1-\alpha_k}$ is the ordered score with rank $r_k = \min\{\lceil (n_c + 1)(1 - \alpha_k) \rceil, n_c\}$. Thus, Δ^L moves the lower boundary downward and Δ^U moves the upper boundary upward. Equation (11) converts the two score distributions into separate shortage-side and accommodation-side protection levels.

3.3. Context-Aware Calibration

Renewable-power forecast errors are condition-dependent, and wind-power forecasting studies have long emphasized the role of lead time and weather regime [6]. Solar-power probabilistic forecasting shows similar dependence on meteorological inputs and post-processing choices [12]. Recent wind-speed and wind-power forecasting studies further highlight the value of probabilistic error modeling under changing forecast conditions [21]. Distribution-learning approaches provide another example of modeling residual uncertainty in renewable-power forecasting [22]. Error distributions vary with renewable type, lead time, output level, weather regime, and prediction difficulty. A single global correction can therefore be too coarse: larger corrections may be required under high weather risk or long lead times, while smaller corrections can preserve sharpness under stable conditions.

CA-ACQR introduces a context variable $\mathbf{z}_t = h(\mathbf{x}_t)$, which is constructed from day-ahead information related to the forecast-error state. In this framework, $\mathbf{z}_t$ includes renewable type, lead-time group, prediction-difficulty regime, weather-risk regime, and output-level bin. The context variables do not change the base model directly; instead, they partition calibration samples so that conformal corrections reflect local error regimes. With context group $g_t = G(\mathbf{z}_t)$, the group-specific corrections and calibrated boundaries are

$$\begin{aligned} \Delta_g^L &= Q_{1-\alpha_L}(\{s_t^L : g_t = g, t \in \mathcal{C}\}), \quad \Delta_g^U = Q_{1-\alpha_U}(\{s_t^U : g_t = g, t \in \mathcal{C}\}), \\ \hat{L}_t &= \hat{q}_{\alpha_L, t}^0 - \Delta_{g_t}^L, \quad \hat{U}_t = \hat{q}_{1-\alpha_U, t}^0 + \Delta_{g_t}^U. \end{aligned} \quad (12)$$

In Equation (12), $h(\cdot)$ extracts the context vector $\mathbf{z}_t$ from day-ahead features, $G(\cdot)$ maps that vector to group g_t , and $\Delta_{g_t}^L$ and $\Delta_{g_t}^U$ are the corrections selected for sample t . The outputs $\hat{L}_t$ and $\hat{U}_t$ are therefore conditioned on the forecast-error regime represented by the context group.

Overly fine context partitioning can make empirical quantiles unstable. Therefore, a minimum calibration sample count $n_{\min}$ is used, and the correction reverts to the global estimate when a group has too few samples:

$$\Delta_g^k = \begin{cases} Q_{1-\alpha_k}(\{s_t^k : g_t = g, t \in \mathcal{C}\}), & n_g \geq n_{\min}, \\ \Delta_{\text{global}}^k, & n_g < n_{\min}, \end{cases} \quad k \in \{L, U\}. \quad (13)$$

Here, n_g is the number of calibration samples assigned to group g , $n_{\min}$ is the minimum admissible group size, and Δ_{global}^k is the side- k correction obtained from all calibration samples. Equation (13) uses a group-specific correction when $n_g \geq n_{\min}$ and the global correction otherwise. For an admissible group, the ordered-score rank is $r_{g,k} = \min\{[(n_g + 1)(1 - \alpha_k)], n_g\}$. The resulting guarantee is marginal within fixed calibration partitions under exchangeability; context grouping improves local adaptation but should not be interpreted as exact conditional coverage for every context state.

3.4. Profile-Dependent Tail-Risk Allocation

In a directional risk interval, the lower and upper sides are operating boundaries rather than purely statistical tails. Under a fixed coverage target, the total tail-risk budget is allocated as

$$\begin{aligned} \alpha_L + \alpha_U &= \alpha, & 0 < \lambda < 1, \\ \lambda &= \frac{\alpha_L}{\alpha}, \\ \alpha_L &= \lambda\alpha, \\ \alpha_U &= (1 - \lambda)\alpha. \end{aligned} \quad (14)$$

The allocation parameter $\lambda = \alpha_L / \alpha$ is the fraction of the fixed total miscoverage budget assigned to the lower tail. Equation (14) keeps $\alpha_L + \alpha_U = \alpha$ and the nominal coverage $1 - \alpha$ unchanged while reallocating risk tolerance between the two sides. A supply-priority profile sets $\lambda < 0.5$ to strengthen the lower boundary, a balanced profile sets $\lambda = 0.5$, and an accommodation-priority profile sets $\lambda > 0.5$ to strengthen the upper boundary. Table 2 summarizes these three profiles.

Table 2. Operating profiles and tail-risk allocation.

Profile	Operating Condition	Tail-Risk Allocation	Expected Effect
Supply-priority	Upward flexibility pressure or supply-adequacy concern	$\alpha_L < \alpha_U$	Lower shortage-side violation
Balanced	Normal operating condition	$\alpha_L = \alpha_U = \alpha/2$	Balanced directional protection
Accommodation-priority	Downward flexibility pressure or renewable accommodation concern	$\alpha_L > \alpha_U$	Lower accommodation-side violation

After an operating profile is selected, CA-ACQR computes $\Delta_{g_t}^L(\alpha_L)$ and $\Delta_{g_t}^U(\alpha_U)$ and forms

$$\begin{aligned} \hat{L}_t &= \hat{q}_{\alpha_L, t}^0 - \Delta_{g_t}^L(\alpha_L), \\ \hat{U}_t &= \hat{q}_{1-\alpha_U, t}^0 + \Delta_{g_t}^U(\alpha_U). \end{aligned} \quad (15)$$

Equation (15) is the final CA-ACQR output: each boundary depends on the base quantile, the context group g_t , and the profile-specific tail probability. Because the initial lower quantile does not exceed the initial upper quantile and both corrections are nonnegative, the calibrated boundaries preserve $\hat{L}_t \leq \hat{U}_t$. The profile design converts a single statistical coverage requirement into a directional risk-control problem aligned with day-ahead operation.

ChatGPT-5.5 was used to assist with drafting and language refinement of selected explanatory passages, which were subsequently reviewed and edited by the authors.

4. Case Study

4.1. Description of Experimental Data

The method is evaluated using a regional wind and PV power dataset. The target variable is renewable output normalized by installed capacity, which makes interval metrics comparable across station sizes and renewable types. The day-ahead input features include historical renewable output, NWP-derived variables, time features, renewable type, and lead-time indicators. Weather-related features represent the influence of wind speed, irradiance, temperature, cloud cover, and other external drivers on renewable variability. Time features represent diurnal cycles, seasonal patterns, and PV night-time zero-output behavior.

The samples are split chronologically into training, calibration, and test sets. The training set fits the base quantile model, the calibration set estimates conformal corrections, and the test set evaluates the final intervals. All compared methods use the same data split and input feature set.

4.2. Description of Case Studies

The evaluation uses a controlled comparison design. All methods share the same chronological training, calibration, and test split; the same base quantile forecasts; the same nominal 90% coverage target; and the same test samples. Directional stress costs are also evaluated with the common weights $c_L = c_U = 1$. Consequently, performance differences arise from the calibration mechanism rather than from data partitions, base forecasts, test composition, or evaluation weights. The three operating profiles preserve the total miscoverage rate $\alpha = 0.10$ and differ only in how that budget is allocated between the lower and upper tails.

Four methods are selected as component benchmarks: base quantile regression (Base-QR), global CQR, asymmetric conformal quantile regression (ACQR), and context-aware conformal quantile regression (CA-CQR). Base-QR evaluates the uncalibrated interval, CQR isolates the effect of a shared global correction, ACQR isolates side-specific calibration, and CA-CQR isolates context grouping. CA-ACQR combines context grouping with asymmetric calibration and profile-dependent tail-risk allocation.

Base-QR directly outputs renewable-power quantiles and is used to evaluate raw interval performance before conformal calibration. CQR applies the global correction in Equation (9). ACQR estimates the two side-specific corrections in Equation (11), whereas CA-CQR incorporates context grouping into a shared interval adjustment. The proposed CA-ACQR jointly integrates side-specific scores, context-aware corrections, and operating-profile-dependent tail-risk allocation, allowing its boundaries to reflect local forecast-error regimes and directional operating preferences.

In practical day-ahead operation, different system states may imply different risk preferences, even when the same nominal coverage target is required. To represent this feature, three operating profiles are considered under a fixed total coverage level. The supply-priority profile reduces the lower-tail risk level to strengthen protection against renewable shortfall and supply-adequacy pressure. The balanced profile adopts symmetric tail-risk allocation and is used as the reference setting for normal operating conditions. The accommodation-priority profile reduces the upper-tail risk level to strengthen protection against renewable surplus, downward-regulation scarcity, and accommodation pressure. Since the three profiles share the same total coverage target, their differences mainly reflect how the same tail-risk budget is redistributed between the two interval sides.

The context variables used by CA-CQR and CA-ACQR are summarized in Table 3. The prediction-difficulty proxy is constructed from pre-target indicators of short-term renewable variability and NWP-derived ramp intensity, while the weather-risk regime is derived from the variability in NWP-based wind speed, irradiance, temperature, and cloud-cover variables. Both variables are discretized according to the tertiles of the calibration set. The lead-time group divides the 96 day-ahead 15 min targets into four 6 h blocks, and the output-level bin is determined from the uncalibrated median forecast and is also discretized by calibration-set tertiles. To prevent unstable empirical quantile estimates caused by overly sparse context groups, the minimum context-group size is set to $n_{\min} = 200$ station-time calibration samples; groups with fewer samples fall back to the global correction.

Table 3. Context variables used for CA-ACQR calibration.

Context Key	Definition	Number of Groups	Purpose
Renewable type	Wind or PV station label	2	Separates resource-dependent error regimes
Lead-time group	15 min day-ahead targets aggregated into four 6 h blocks	4	Captures horizon-dependent uncertainty
Prediction-difficulty regime	Pre-target variability and NWP-derived ramp-intensity proxy, split by calibration-set tertiles	3	Distinguishes low-, medium-, and high-difficulty samples
Weather-risk regime	NWP-derived wind-speed, irradiance, temperature, and cloud-cover variability, split by calibration-set tertiles	3	Represents weather-driven forecast-risk conditions
Output-level bin	Uncalibrated median forecast level, split by calibration-set tertiles	3	Accounts for output-dependent error behavior

For the stress-cost comparison, all strategies are evaluated on the same test samples with fixed symmetric weights, i.e., $c_L = c_U = 1$. The fixed-reserve strategy uses a constant symmetric margin determined from the empirical 90% interval width of the calibration set, whereas the historical-quantile strategy uses empirical lower and upper renewable-output quantiles estimated from the calibration split. Vanilla CQR denotes the global CQR interval reported in Table 4, while CA-ACQR denotes the proposed profile-dependent context-aware asymmetric interval. The mean stress cost is averaged over all station-time test samples, and CVaR₉₅ is computed from the upper 5% tail of the same test-sample cost distribution.

Table 4. Overall performance comparison of prediction intervals on the test set.

Method	PICP (%)	PINAW	VR ^L (%)	VR ^U (%)	Mean Stress Cost	CVaR ₉₅
Base-QR	91.11	0.459	4.52	4.37	0.332	0.816
CQR	91.11	0.459	4.52	4.37	0.332	0.816
CA-CQR	91.64	0.412	4.11	4.25	0.333	0.816
ACQR	93.45	0.487	4.52	2.03	0.356	0.824
CA-ACQR	94.07	0.367	4.49	1.44	0.360	0.836

4.3. Analysis of Experimental Results

Table 4 and Figure 2 summarize the overall calibration performance of the compared interval methods. Base-QR and CQR obtain the same PICP of 91.11%, with shortage-side and accommodation-side violation rates of 4.52% and 4.37%, respectively, because

the global CQR correction is inactive when the raw $q_{0.05}$ – $q_{0.95}$ interval already satisfies the calibration requirement. This result is therefore used as the reference performance before introducing context-aware and asymmetric calibration. As shown by the coverage-gap panel, all methods exceed the nominal 90% level, while the reliability–sharpness panel further reveals how the methods differ in terms of coverage improvement and interval width.

Introducing context-aware calibration changes the interval structure in a more adaptive manner. CA-CQR increases PICP from 91.11% to 91.64%, reduces PINAW from 0.459 to 0.412, and decreases the shortage-side and accommodation-side violation rates to 4.11% and 4.25%, respectively. These results indicate that context grouping does not simply widen the prediction interval to increase coverage; instead, it adjusts the correction according to renewable type, prediction difficulty, weather-risk regime, and lead time, thereby improving reliability while reducing the average interval width. Asymmetric calibration has a more pronounced effect on directional violation control. ACQR increases PICP to 93.45% and reduces the accommodation-side violation rate from 4.37% to 2.03%, while the shortage-side violation rate remains at 4.52%. By further combining context-aware adjustment with asymmetric side-specific calibration, CA-ACQR increases PICP to 94.07% and reduces the accommodation-side violation rate to 1.44%, corresponding to a reduction of 2.93 percentage points relative to Base-QR. Meanwhile, CA-ACQR achieves the lowest PINAW among the compared methods, which shows that the improved coverage and directional protection are not obtained by a simple expansion of the interval. This reliability and sharpness advantage, however, should be interpreted together with the stress-cost results because the proposed method is designed for selectable risk posture rather than for uniformly minimizing all cost-related indicators.

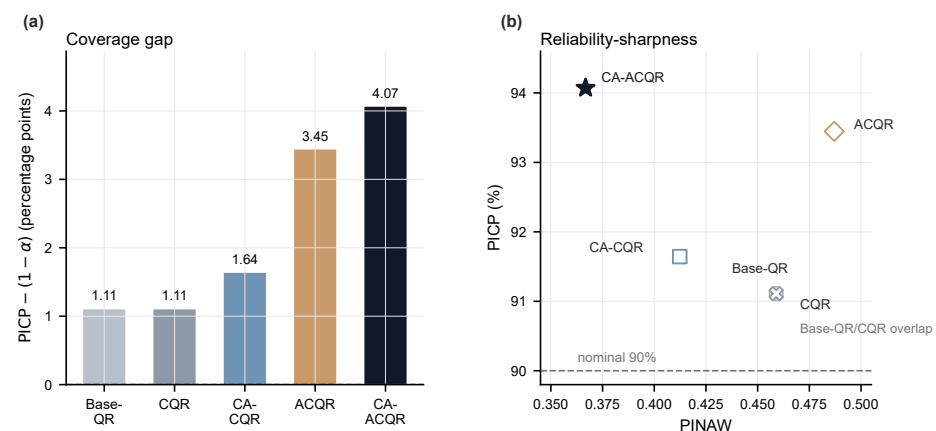


Figure 2. Reliability–sharpness comparison of the interval calibration methods. Panel (a) reports the coverage gap relative to the nominal 90% level, $\text{PICP} - (1 - \alpha)$, in percentage points. Panel (b) shows the PICP–PINAW trade-off and explicitly marks the overlap between Base-QR and CQR.

Figure 3 illustrates how CA-ACQR redistributes directional risk through the tail-allocation parameter $\lambda = \alpha_L / \alpha$. In the continuous allocation panel, increasing λ transfers protection from the lower boundary to the upper boundary: VR^L increases from 0.86% to 4.02%, whereas VR^U decreases from 4.31% to 0.69%. This trend is consistent with the proposed boundary interpretation, because reducing α_L strengthens shortage-side protection, while reducing α_U strengthens accommodation-side protection.

The representative operating profiles further show the practical meaning of this continuous allocation mechanism. Under the supply-priority profile, the shortage-side violation rate is reduced to 0.86%, which is 2.83 percentage points lower than the balanced value of 3.69%. Under the accommodation-priority profile, the accommodation-side violation

rate is reduced to 0.69%, which is 2.96 percentage points lower than the balanced value of 3.65%. These reductions are accompanied by a shift of risk to the opposite side, as reflected by the higher accommodation-side violation under supply priority and the higher shortage-side violation under accommodation priority. The width and stress-cost panels show that PINAW and mean stress cost remain relatively low around the supply-priority to balanced region, but increase under the strongest accommodation-priority setting. Therefore, operating-profile selection should be understood as choosing a point on a trade-off surface among supply-adequacy risk, accommodation risk, interval width, and stress cost, rather than as selecting a universally dominant interval.

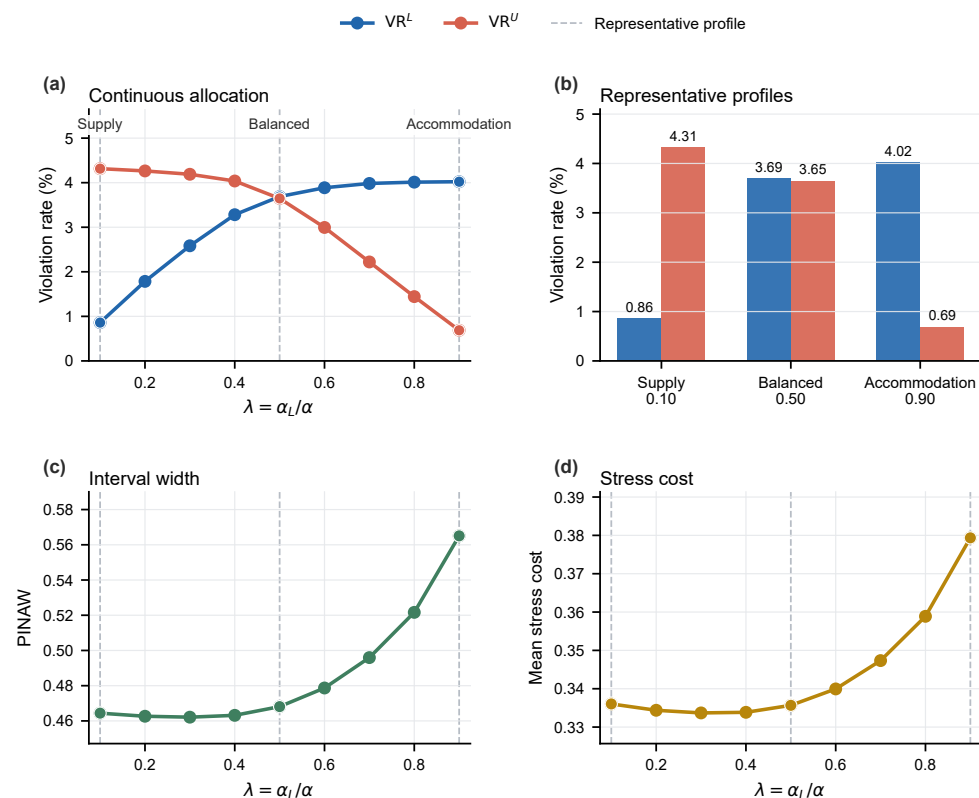


Figure 3. Operating profiles and tail-risk allocation in CA-ACQR: (a) continuous changes in the shortage-side violation rate VR^L and accommodation-side violation rate VR^U with the tail-allocation parameter $\lambda = \alpha_L / \alpha$, where the vertical dashed lines mark the representative profiles; (b) directional violation rates under the supply-priority, balanced, and accommodation-priority profiles, where the blue bars denote VR^L and the red bars denote VR^U ; (c) variation in PINAW with λ ; and (d) variation in mean stress cost with λ .

The stress-cost trade-off in Figure 4 provides a complementary view of the operating implications of directional protection. The two conventional strategies, fixed reserve and historical quantile, show much larger tail stress-cost values, indicating that static margins or historical quantiles may be insufficient for controlling difficult test periods under the adopted normalized penalty assumptions. Vanilla CQR and CA-ACQR are located in the lower-cost region of the trade-off space, but CA-ACQR is not the minimum point for every cost metric. Its mean stress cost is 0.360, higher than that of Base-QR/CQR (0.332), and its $CVaR_{95}$ is 0.836, slightly higher than that of Base-QR/CQR (0.816). This result indicates that stronger directional boundary protection can reduce the selected violation rate, but the corresponding boundary shifts and fixed normalized penalty weights may increase aggregate stress-cost indicators. Therefore, CA-ACQR should be interpreted as a risk-posture selection mechanism that supports directional protection under a fixed coverage target, rather than as a method that uniformly minimizes all average and tail cost measures.

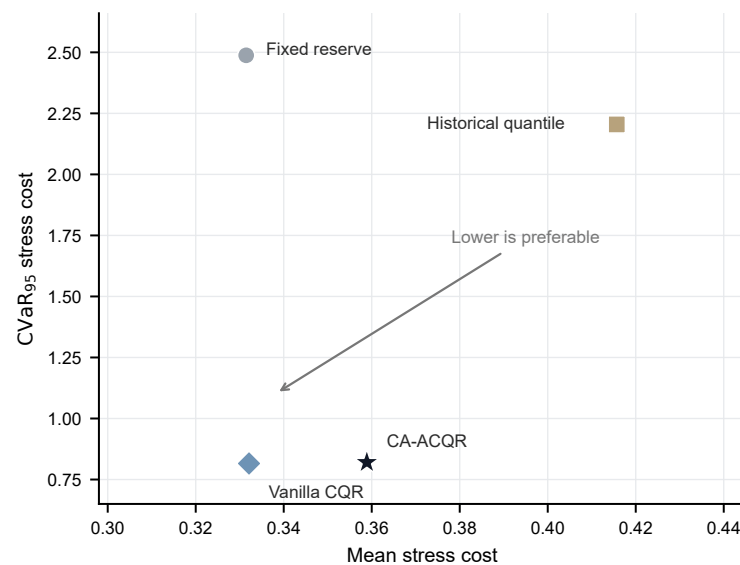


Figure 4. Trade-off between mean stress cost and CVaR₉₅ stress cost across operating-risk strategies. Each point represents one strategy evaluated on the same test samples; movement toward the lower-left direction indicates lower average and tail stress-cost values under the adopted normalized penalty assumptions.

Figure 5 decomposes the tail-allocation response by renewable type. Wind and PV stations exhibit the same qualitative profile-control pattern: as λ increases from the supply-priority side to the accommodation-priority side, lower-side protection becomes weaker and VR^L increases, while upper-side protection becomes stronger and VR^U decreases. The shared y-axis makes the difference between the two resource types more explicit. Wind stations show a higher accommodation-side violation rate under supply-priority allocation and a larger reduction as λ increases, whereas PV stations exhibit a lower but more gradual accommodation-side response. These differences in curve levels and slopes indicate resource-dependent error regimes, supporting the use of context-aware calibration rather than a single global correction.

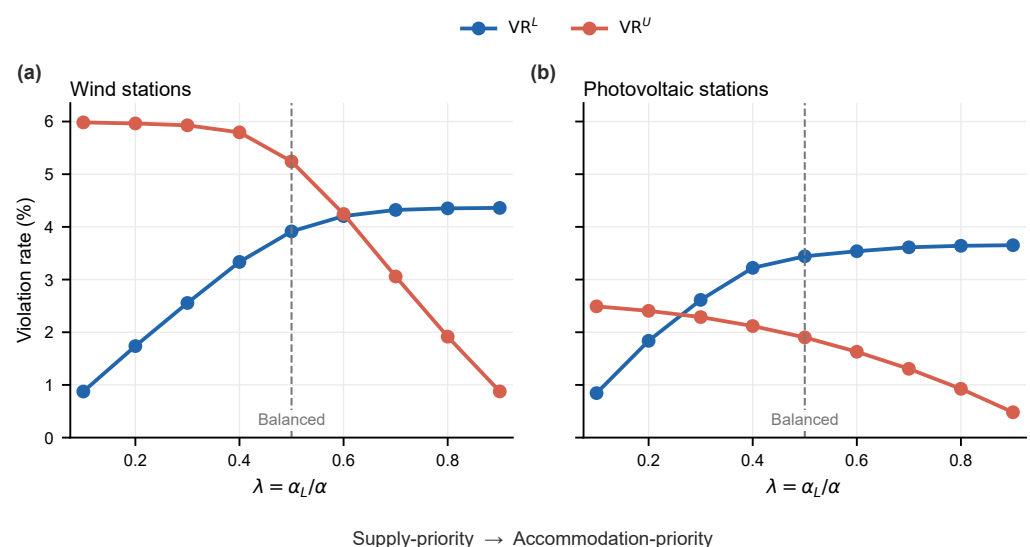


Figure 5. Directional violation response to tail-risk allocation for wind and photovoltaic stations: (a) violation-rate responses for wind stations; and (b) violation-rate responses for photovoltaic stations. In both panels, the blue curve denotes the shortage-side violation rate VR^L , the red curve denotes the accommodation-side violation rate VR^U , and the vertical dashed line marks the balanced allocation $\lambda = 0.5$.

4.4. Discussion

The experimental results demonstrate that CA-ACQR provides an effective mechanism for converting prediction intervals into directional operating risk boundaries. Traditional probabilistic interval methods mainly evaluate overall coverage and width, which are necessary for assessing reliability and sharpness but are insufficient for identifying whether violations occur on the shortage side or on the accommodation side. CQR improves the conformal-calibration baseline by applying a global correction, but it does not distinguish the two directions of operational risk. ACQR separates lower- and upper-side violations but still relies on global side-specific corrections, whereas CA-CQR accounts for forecast-error heterogeneity through context grouping but does not allocate tail risk according to operating profiles. CA-ACQR combines these elements and produces calibrated interval boundaries that depend jointly on local forecast-error regimes and operational risk preferences.

From an operational perspective, the three profiles correspond to different day-ahead risk postures. Supply-priority intervals are more suitable when supply pressure, upward flexibility scarcity, or reserve adequacy is the main concern, whereas accommodation-priority intervals are more suitable when renewable surplus, downward-regulation scarcity, or accommodation pressure becomes dominant. Balanced intervals provide a reference posture for regular operating conditions. These profiles do not imply a fixed performance ranking; instead, they provide selectable risk postures under the same total coverage target, allowing operators to choose the interval boundary configuration that is more consistent with the prevailing operational concern.

It should also be emphasized that CA-ACQR provides day-ahead probabilistic risk boundaries rather than replacing unit commitment, economic dispatch, or security-checking models. Its role is to generate directional uncertainty inputs that can support downstream reserve scheduling, dispatch decision-making, and renewable accommodation assessment. The detailed operating-cost impacts of the proposed intervals should therefore be evaluated after embedding these calibrated boundaries into downstream optimization or security-analysis models.

5. Conclusions

This paper proposed CA-ACQR, a context-aware asymmetric conformal calibration framework for constructing directional renewable-power risk intervals. The method interprets the lower boundary as a shortage-side supply-adequacy risk boundary and the upper boundary as an accommodation-side renewable-accommodation risk boundary. It uses side-specific conformal scores, context-aware corrections, and profile-dependent tail-risk allocation to adjust the two interval boundaries under supply-priority, balanced, and accommodation-priority operation.

The regional wind and PV case study shows that CA-ACQR improves interval reliability and directional violation control while exposing measurable stress-cost trade-offs. In the component comparison, CA-ACQR increases PICP to 94.07% and reduces accommodation-side violation from 4.37% to 1.44%. Along the operating-profile frontier, supply-priority operation reduces shortage-side violation to 0.86%, whereas the strongest accommodation-priority setting reduces accommodation-side violation to 0.69%. These results indicate that the proposed directional intervals can support day-ahead risk screening, reserve awareness, renewable-accommodation assessment, and operating-profile selection, while downstream operational optimization remains necessary for detailed dispatch decisions.

Author Contributions: Conceptualization, P.H. and R.L.; methodology and formal analysis, R.L. and Y.W.; software and visualization, Y.W.; data curation and validation, P.H., Y.L. and X.Y.; resources, J.Z. and P.H.; writing—original draft preparation, Y.W., Y.L. and X.Y.; writing—review and editing, J.Z.

and R.L.; supervision and project administration, P.H. and R.L.; funding acquisition, P.H. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of State Grid Shanxi Electric Power Co., Ltd., grant number 52053026000M.

Data Availability Statement: The data that support the findings of this study are subject to restrictions from the data provider and are not publicly available. Aggregated or processed results needed to reproduce the main figures can be made available from the corresponding author upon reasonable request, subject to data-use restrictions.

Acknowledgments: The authors acknowledge the support of State Grid Shanxi Electric Power Co., Ltd. During the preparation and revision of this manuscript, the authors used ChatGPT-5.5 (OpenAI, accessed June–July 2026) for drafting and language refinement. The authors reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: Authors Peng Han and Yu Liu were employed by the company Electric Power Dispatching and Control Center, State Grid Shanxi Electric Power Co., Ltd. Author Jun Zhao was employed by the company Electric Power Research Institute, State Grid Shanxi Electric Power Co., Ltd. Author Xuehai Yu was employed by the company Datong Power Supply Company, State Grid Shanxi Electric Power Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare that this study received funding from State Grid Shanxi Electric Power Co., Ltd. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. Tuohy, A.; Meibom, P.; Denny, E.; O'Malley, M. Unit Commitment for Systems with Significant Wind Penetration. *IEEE Trans. Power Syst.* **2009**, *24*, 592–601. [\[CrossRef\]](#)
2. Ortega-Vazquez, M.A.; Kirschen, D.S. Estimating the Spinning Reserve Requirements in Systems with Significant Wind Power Generation Penetration. *IEEE Trans. Power Syst.* **2009**, *24*, 114–124. [\[CrossRef\]](#)
3. Hirth, L.; Ziegenhagen, I. Balancing Power and Variable Renewables: Three Links. *Renew. Sustain. Energy Rev.* **2015**, *50*, 1035–1051. [\[CrossRef\]](#)
4. Bird, L.; Cochran, J.; Wang, X. *Wind and Solar Energy Curtailment: Experience and Practices in the United States*; Technical Report NREL/TP-6A20-60983; National Renewable Energy Laboratory: Golden, CO, USA, 2014. [\[CrossRef\]](#)
5. Hong, T.; Pinson, P.; Wang, Y.; Weron, R.; Yang, D.; Zareipour, H. Energy Forecasting: A Review and Outlook. *IEEE Open Access J. Power Energy* **2020**, *7*, 376–388. [\[CrossRef\]](#)
6. Pinson, P. Wind Energy: Forecasting Challenges for Its Operational Management. *Stat. Sci.* **2013**, *28*, 564–585. [\[CrossRef\]](#)
7. Gneiting, T.; Katzfuss, M. Probabilistic Forecasting. *Annu. Rev. Stat. Its Appl.* **2014**, *1*, 125–151. [\[CrossRef\]](#)
8. Koenker, R.; Bassett, G. Regression Quantiles. *Econometrica* **1978**, *46*, 33–50. [\[CrossRef\]](#)
9. Wan, C.; Lin, J.; Wang, J.; Song, Y.; Dong, Z.Y. Direct Quantile Regression for Nonparametric Probabilistic Forecasting of Wind Power Generation. *IEEE Trans. Power Syst.* **2017**, *32*, 2767–2778. [\[CrossRef\]](#)
10. Wan, C.; Xu, Z.; Pinson, P.; Dong, Z.Y.; Wong, K.P. Optimal Prediction Intervals of Wind Power Generation. *IEEE Trans. Power Syst.* **2014**, *29*, 1166–1174. [\[CrossRef\]](#)
11. Wan, C.; Wang, J.; Lin, J.; Song, Y.; Dong, Z.Y. Nonparametric Prediction Intervals of Wind Power via Linear Programming. *IEEE Trans. Power Syst.* **2018**, *33*, 1074–1076. [\[CrossRef\]](#)
12. Doubleday, K.; Jascourt, S.; Kleiber, W.; Hodge, B.M. Probabilistic Solar Power Forecasting Using Bayesian Model Averaging. *IEEE Trans. Sustain. Energy* **2021**, *12*, 325–337. [\[CrossRef\]](#)
13. Shafer, G.; Vovk, V. A Tutorial on Conformal Prediction. *J. Mach. Learn. Res.* **2008**, *9*, 371–421. [\[CrossRef\]](#)
14. Lei, J.; G'Sell, M.; Rinaldo, A.; Tibshirani, R.J.; Wasserman, L. Distribution-Free Predictive Inference for Regression. *J. Am. Stat. Assoc.* **2018**, *113*, 1094–1111. [\[CrossRef\]](#)
15. Romano, Y.; Patterson, E.; Candès, E.J. Conformalized Quantile Regression. *arXiv* **2019**. [\[CrossRef\]](#)
16. Barber, R.F.; Candès, E.J.; Ramdas, A.; Tibshirani, R.J. Predictive Inference with the Jackknife+. *Ann. Stat.* **2021**, *49*, 486–507. [\[CrossRef\]](#)
17. Bunodiare, A.; Lee, H.S. Renewable Energy Curtailment: Prediction Using a Logic-Based Forecasting Method and Mitigation Measures in Kyushu, Japan. *Energies* **2020**, *13*, 4703. [\[CrossRef\]](#)

18. Gneiting, T.; Raftery, A.E. Strictly Proper Scoring Rules, Prediction, and Estimation. *J. Am. Stat. Assoc.* **2007**, *102*, 359–378. [[CrossRef](#)]
19. Schröder, T.; Kuckshinrichs, W. Value of Lost Load: An Efficient Economic Indicator for Power Supply Security? A Literature Review. *Front. Energy Res.* **2015**, *3*, 55. [[CrossRef](#)]
20. Sullivan, M.; Schellenberg, J.; Blundell, M. *Updated Value of Service Reliability Estimates for Electric Utility Customers in the United States*; Technical Report LBNL-6941E; Lawrence Berkeley National Laboratory: Berkeley, CA, USA, 2015. [[CrossRef](#)]
21. Chaouch, M. Probabilistic Wind Speed Forecasting for Wind Turbine Allocation in the Power Grid. *Energies* **2023**, *16*, 7615. [[CrossRef](#)]
22. Chen, F.; Gao, L. Learning Residual Distributions with Diffusion Models for Probabilistic Wind Power Forecasting. *Energies* **2025**, *18*, 4226. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.