Electrical Power and Energy Systems 157 (2024) 109889

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Modeling wind energy imbalance risk in medium-term generation planning models: A methodological proposal for large scale applications

Geovanny Marulanda [a,*], Antonio Bello [a], Javier Reneses [b]

[a] Institute for Research in Technology (IIT), ICAI School of Engineering, Comillas Pontifical University, Calle de Santa Cruz de Marcenado, 26, Madrid, 28015, Madrid, Spain

[b] Simulart Energy, Foronda, 6, Madrid, 28034, Madrid, Spain

ARTICLE INFO

Keywords:
Generation planning
Imbalance risk
Real-size power systems modeling
Temporal aggregation
Wind energy forecasting

ABSTRACT

Wind power is one of the most important renewable energy sources worldwide. However, integrating these technologies in electricity markets faces challenges due to the intermittent wind speed, resulting in potential imbalances within power grids and subsequent inefficiencies. Traditional approaches for modeling the uncertainty simulate fixed wind power values in operational and planning models, which work well for the short term but are inadequate for medium-term decision-making due to lower forecasting accuracy and the inability to differentiate the imbalance risk faced in each simulated value. This paper proposes a novel methodology for incorporating wind energy imbalance risk into medium-term generation planning models. For this, we first simulate realistic wind power scenarios and perform a fitting distribution process to characterize the hourly wind power behavior. Then, we model the imbalance risk using quadratic-uncertainty cost functions for each simulated value. Subsequently, we conduct a temporal aggregation to reduce the computational burden, preserving the conceptual framework of the imbalance risk modeling. Finally, we propose a medium-term generation planning model that includes the imbalance risk model and the temporal aggregation strategy. Our approach is evaluated in a real-size case study formed by Spain, Portugal, and France's electricity markets, demonstrating accurate wind power imbalance risk modeling and lowered planning costs.

1. Introduction

Wind power is one of the most important renewable energy sources worldwide. In fact, it is the second renewable energy source after hydroelectric power and above solar generation [1]. Recently, wind energy sources have seen significant growth in power networks, owing to environmental concerns such as reducing greenhouse gas emissions and technological advancements leading to lower investment costs [2]. Furthermore, yearly capacity increases for onshore and offshore wind power have been predicted. On the one hand, the onshore capacity is expected to reach more than 200 GW globally by 2050, more than 24 times the amount of wind power installed in 2018 [3]. On the other hand, it is predicted that the global cumulative of offshore technologies will have increased sevenfold to 215 GW or more by 2030 [4].

However, wind speed's intermittence impedes its widespread integration into power grids because it could lead to uncertainty and fluctuation in electricity supply [5]. In fact, energy oscillation needs extra capacity in the power grid, lowering the economy of the system and its stability [6]. In addition, wind-power deviations result in increased requirements for ancillary services and spinning reserves, leading to higher overall operating costs [7]. One of the primary challenges faced by system operators and market agents is to effectively utilize this variable resource while maintaining a balance between consumption and generation, ensuring high system reliability. Nevertheless, the participation of wind producers in energy markets is hindered by the uncertainty surrounding their power generation. Imbalance costs can be considerably high, prompting wind producers to offer lower output than expected in order to mitigate these costs [8,9]. This strategy could result in reduced profits and inefficiency. In the same way, planning for wind energy in medium- and long-term scheduling are extremely challenging because of the erratic and intermittent nature of wind power supply [10]. In particular, it is difficult to accurately estimate the effects of wind power integration on electricity markets, which increases the risk exposure of the market agents and hinders the decision-making process in planning-related issues.

In this regard, numerous techniques have emerged to assess the impacts of the stochastic nature of wind in electricity markets. In [11], a risk-seeking stochastic optimization model for trading wind power in electricity markets is proposed. The wind power scenarios are

* Corresponding author.
E-mail addresses: gmarulanda@comillas.edu (G. Marulanda), abello@comillas.edu (A. Bello), javier.reneses@simulart.es (J. Reneses).

https://doi.org/10.1016/j.ijepes.2024.109889
Received 28 June 2023; Received in revised form 29 December 2023; Accepted 19 February 2024
Available online 22 February 2024

generated using a seasonal autoregressive integrated moving average (SARIMA) approach. Subsequently, these scenarios are integrated into a risk-seeking stochastic optimization model that considers the real-time wind power deviations. In [12], a stochastic programming model is proposed for the integrated operation of hydro and wind resources. The authors utilize a martingale model of forecast evolution (MMFE) to simulate the wind power uncertainty. This approach effectively models the evolution of forecast uncertainty associated with wind power, enabling the generation of long-range synthetic forecast scenarios. A comprehensive review of the fundamental steps involved in stochastic optimization and uncertainty modeling within renewable energy applications context is offered in [13].

The most common approaches for tackling uncertainty in power systems are probabilistic techniques [14]. In these approaches, the probability density functions (PDF) of the uncertain variables as wind, are calculated to specify the chance to which one observation falls within a particular range of values. Once the PDF is obtained, three probabilistic strategies are typically followed to model the effect of the uncertain variable on the electricity market: Sampling-based simulations, scenario-based approaches, and point estimation methods [6, 15]. In sampling-based simulations, discrete scenarios of the uncertain variable are taken from the PDF. After that, simulations on each sample are conducted to characterize the impact of the uncertainty on the market. As more samples are simulated, more precise results are expected. In contrast, the PDF curve is subdivided into uncertain intervals in the scenario-based approaches. Then, representative scenarios are selected considering the variable's probability of falling in each interval. Reduction scenario techniques are usually included to diminish the computing burden. Finally, the point estimation methods use the moments of the uncertain input parameters to build the PDF of the market output to be analyzed [15]. A detailed review of different modeling techniques for the uncertainty applied to energy systems can be found in [14,16–18].

Most of the approaches mentioned above characterize wind's uncertain impact on the electricity market, simulating fixed values of wind power in an operational or planning model. This approach works well for the short-term but could not be the best approach for medium- and long-term generation planning models, mainly for two reasons. On the one hand, the wind is one of the most challenging meteorological data to forecast. At present, it is impossible to predict wind changes with sufficient accuracy for longer horizons than a few days [19,20]. In this way, simulating fixed values could be an extreme measure for scenarios that could be highly inaccurate, like those located on distribution tails. This approach could lead to expensive planning scenarios. On the other hand, it could be unrealistic because this representation does not model the balancing actions through reserves management that system operators would take in the short term [21]. Although some proposed strategies in the literature allow flexibility ranges [22], they do not differentiate the imbalance risk faced in each possible realization. In this context, an appropriate flexible approach with suitable imbalance risk modeling could help find more realistic solutions and better reserve scheduling at lower planning costs.

Analytical formulations have been proposed to model imbalance risk due to wind resources. These approaches enable the integration of wind uncertainty into optimization models by adjusting the model parameters by suitable values. With this approach, it became possible to simulate diverse aspects of forecasting uncertainty concerning meteorological conditions and predicted power magnitude [8]. In [8], the wind power uncertainty dispatch cost (WPUDC) is introduced as the expectation of the operational cost for surplus incurred by wind power forecasting. The WPUDC is a quadratic function obtained from a linear piecewise approach of the CDF of wind power constituted of three segments. The authors show that the linear approximation accurately matches the dispatched wind power in all distribution intervals. In [7], a measure to quantify the probabilistic risk of failing the contracted dispatch is developed using the same linear approach on the wind power CDF. It is concluded that multiple quadratic cost functions better capture the wind uncertainties, differentiate the seasonal behavior, and improve the grid's efficiency. In [23,24], the concept of WPUDC has been extended to model the uncertainty cost of renewable resources in Cournot's oligopolistic games of energy trading. Results show the existence of a unique Nash equilibrium and the suitability of this approach to resolve competition models for local markets. Finally, in [21], a polynomial cost function for chance-constrained economic dispatch problems has been proposed to model parametric and non-parametric distributions of wind power. With this approach, the total operational cost is reduced.

However, the above references primarily focus on short-term planning horizons and small-scale case studies. These settings allow for the availability of accurate expected values, the anticipation of minor variations in actual wind power, and computational constraints that limit the simulation of exhaustive scenarios in planning models are generally absent. In contrast, this paper aims to shed light on some questions that remain open in the literature, including: How to effectively incorporate the high uncertainty of wind power into medium-term generation planning models while adequately accounting for its probabilistic features and its impact on balancing costs? How to reduce the computational burden of large-scale applications when considering numerous scenarios to model wind power uncertainty?

Although these questions have been studied in the literature, these challenges persist, and novel approaches continue to emerge to address them. For instance, recently [25] introduced a model that addresses wind power integration uncertainty in system planning, utilizing a non-parametric kernel density estimation method to fit wind power prediction errors. In contrast, [26] adopts a multi-stage stochastic approach considering the uncertainty of wind power on representative days. Alternatively to these references, we propose a methodology to include wind power uncertainty in medium-term generation planning models through parametric functions. On the other hand, the issue of reducing the computational burden of large-scale applications has been addressed in works such as [27–29], and [30]. These studies present methodologies designed to enhance performance and reduce the computational load when addressing the generation planning problem in large-scale applications. In contrast to these existing approaches, our contribution focuses on addressing computational challenges by introducing temporal aggregation for the representation of the wind power uncertainty.

In this context, the main contributions of our work to the state of the art can be summarized as follows:

1. We introduce a novel methodology for incorporating wind energy imbalance risk into medium-term generation planning models. In contrast to existing literature, our approach involves calculating individual quadratic cost functions for each potential wind resource realization. This calculation depends on its location on the PDF and the maximum permissible deviation between the realization and the actual dispatch. This approach enables us to accurately incorporate the high uncertainty of wind power in the mid-term into generation planning models while considering the imbalance risk and balancing cost faced by each scenario.
2. We introduce the use of uncertainty cost functions aggregated temporally to address computational challenges in large-scale applications. This approach reduces the computational burden by reducing the granularity of the problem and relaxing the constraints associated with the balance between generation and demand. Moreover, after temporal aggregation, our approach preserves the main features identified hourly for the uncertainty, imbalance risk, and balancing cost modeling.

The rest of the paper is structured as follows. Section 2 introduces the proposed methodology and describes the medium-term generation planning model. Section 3 shows the results obtained for the study case. Finally, the main findings of this work are shown in Section 4.

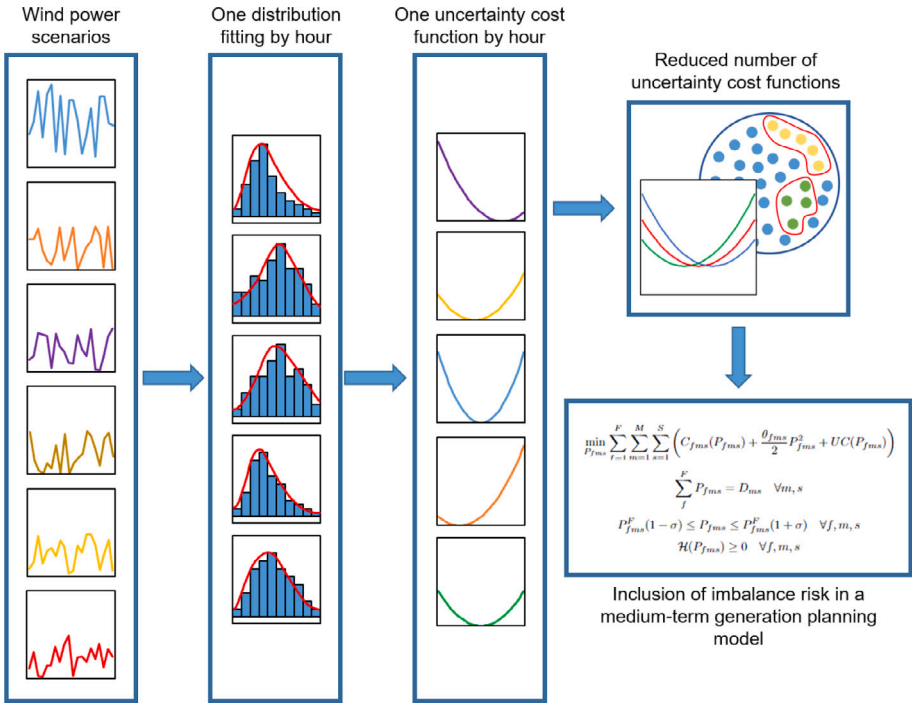

Fig. 1. Schematic representation of the proposed methodology.

2. Methodology

Fig. 1 provides an overview of the methodology proposed in this paper. Firstly, several realistic wind power scenarios are simulated for each hour of the midterm horizon, using the methodology proposed in [31]. Next, a fitting distribution process is performed to represent the observed values of each hour through hourly CDFs. These CDFs are then utilized as a basis for modeling imbalance risk through quadratic-uncertainty cost functions. Subsequently, a temporal aggregation has been conducted to reduce the complexity and computational burden of the medium-term generation planning model. Moreover, one temporal aggregation strategy is proposed to include the uncertainty cost functions in large-size problems. Finally, a medium-term generation planning model that incorporates these functions is suggested. The subsequent sections provide detailed explanations of the methodology presented in this paper.

2.1. Scenarios generation

This paper follows the methodology proposed in [31] to generate realistic wind power scenarios. This approach allows for including the spatial–temporal dependencies of wind power among multiple interconnected electricity markets with an hourly resolution. For this, a set of scenarios is generated representing the total wind power production for each geographical area. Unlike approaches that focus on detailed representations of individual wind turbines and farms like [32,33], this aggregated methodology offers valuable insights into the impact of wind behavior across these markets. It considers factors such as network status and operational strategies that local methods may not adequately capture. These factors play an essential role in mid-term planning models, where understanding the behavior of the interconnected system is crucial to modeling the integrated market.

In this context, multiple-time series of wind power are simulated for each area using the value of wind power utilization instead of real power. This approach removes dependency and trends from the historical data due to the wind power installed capacity, leading to a better representation of spatial–temporal correlations among areas. In short, this methodology is divided into three main steps. First, a time series decomposition of historical data is performed to capture the trends and seasonality of each interconnected market. Second, Seasonal Auto-Regressive Integrated Moving Average models (SARIMA) are developed to obtain the temporal dependencies of each area. Finally, the simulation of several paths of wind power considering the spatial dependencies is conducted through correlated and multivariate Monte Carlo simulations. This methodology closely simulates the dynamics of the PDF of the historical data at monthly, weekly, daily, and hourly scales, making it suitable for medium-term planning applications. It is important to note that the high accuracy and reliability of the SARIMA model depend on various factors, including data reliability, parameter selection, model assumptions, spatial correlations, and uncertainty processing. Therefore, a thorough evaluation of the quality of the scenarios generated should be conducted, tailored to the specific study case. Detailed measures assessing the quality of the scenarios generated for this paper can be found in [31].

2.2. Distribution fitting

After obtaining wind power paths from the scenario generator, a fitting distribution is proposed to be performed for each hour. To achieve this, the PDF of each hour was obtained and adjusted through maximum likelihood estimation. The adequacy of the fitting was tested using the Kolmogorov–Smirnov (KS) test [34]. The KS test evaluates the null hypothesis that the data from the scenario generator follows a specific distribution against the alternative hypothesis that it does not. In particular, we have considered that the null hypothesis is accepted at the 5% significance level, indicating a high level of accuracy in the selected distribution. In this paper, six distributions proposed in the literature for modeling wind power density functions were considered [35–37]: Normal, Weibull, Gamma, Beta, Logistic, and Kernel distributions. However, other distributions could be considered in our methodology depending on the specific application [35], and a detailed analysis of whether these distributions can accurately describe the data distribution is encouraged to be explored in future works. If multiple distributions fit the data, Akaike's Information Criteria (AIC) was used to compare their relative quality and determine the best fit [38]. In short, through the KS test and the AIC we guarantee that the probability

distribution fitted have been adequately selected and describe the data properly. Finally, the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) measures have been calculated to assess the quality of the fit for each hour as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

where n is the sample size, y_i and $\hat{y}_i$ are the scenario and the fitted value for the scenario i, respectively. The MAE and the RMSE have been widely used due to their simplicity and explainability. In particular, MAE gives less weight to outliers, while RMSE gives a relatively high weight to large errors [39]. In conjunction, the information of both measures gives a sensitivity of the number and magnitude of the errors obtained with the fitting. As will be seen in the next section, fitting the generated scenarios to known distribution functions enables obtaining polynomial approaches around the realizations and accurately modeling the uncertainty cost of wind power.

2.3. Uncertainty cost

Once a PDF has been fitted by each hour, the imbalance risk of each hour is modeled. Usually, an imbalance cost is produced because of the under- or over-estimation of the actual wind power in the forecasting stage. In [8] the concept of wind power uncertainty incremental cost (UIC) is introduced to quantify the extra cost by unit for balancing wind power uncertainty in system operation as:

$$UIC(w_F) = \frac{\partial E[C(\Delta w_F)]}{\partial w_F} \tag{3}$$

where $C(\Delta w_F)$ is the cost for balancing power deviation from the forecasted wind power w_F to the real wind power generation w_R, and $E[C(\Delta w_F)]$ is the expected value of the balancing cost. Likewise, the upward and downward reserve costs are given by Eqs. (4)(a) and (b), respectively:

$$C(\Delta w_F) = \begin{cases} C_r(w_F) = k_r(w_F - w_R), & \text{if } w_F - w_R > 0; \quad \text{(a)} \\ C_p(w_F) = k_p(w_R - w_F), & \text{if } w_R - w_F > 0; \quad \text{(b)} \end{cases} \tag{4}$$

where k_r and k_p are the unit cost of the upward and downward reserves. Notice from (3) that the uncertainty cost proposed by [8] depends on the expected value of the balancing cost. This approach is valid for the short term when an accurate expected value can be found and minor variations for the actual wind power are anticipated. However, the uncertainty is much greater in the medium term. This paper argues that the deviations concerning each possible realization should be considered in the modeling to better capture the imbalance risk at different distribution intervals. Therefore, in contrast to what was proposed by [8], this paper introduces a novel definition of the expected value of the imbalance cost for a realization w_F between w_F^- and w_F^+ by:

$$E[C(\Delta w_F)] = \int_{w_F^-}^{w_F^+} C(\Delta w_F) f(w) dw \tag{5}$$

$$= k_r \int_{w_F^-}^{w_F} (w_F - w) f(w) dw + k_p \int_{w_F}^{w_F^+} (w - w_F) f(w) dw \tag{6}$$

where $f(w)$ is the PDF of the wind power w. In this paper, it is assumed for simplicity that $w_F^- = w_F - \sigma$ and $w_F^+ = w_F + \sigma$, where σ is the deviation interval and also $\sigma \leq w_f$. The parameter σ represents the permissible range of variations for wind energy technologies. These deviations encompass both the difference between the market-bid (forecasted value) and the actual dispatch, as well as the maximum allowable balancing actions through reserves that system operators might need to implement in the short term [21,40]. The acceptable deviation limits can vary based on factors like the technology employed, the country's regulations, and the unique characteristics of the electricity market. Consequently, the appropriate value for σ should be determined in consultation with the specific electricity market in question. Then, replacing (6) in (3) it is obtained that:

$$UIC(w_F) = k_r \cdot F(w)\Big|_{w_F^-}^{w_F} + k_p \cdot F(w)\Big|_{w_F}^{w_F^+} \tag{7}$$

where $F(w)$ is the CDF of w. In contrast to [8], where the entire CDF is approached through a piecewise function of three linear terms, in this paper, we focus on each individual realization and approximate the CDF using a straight line in the vicinity of that realization. By adopting this approach, we achieve a more refined and localized representation of the CDF, offering improved accuracy and flexibility compared to the previous method. As the proposed fitting process gives differentiable functions, we approach the CDF around each realization through the first term of Taylor's series. It is a common procedure applied in engineering problems, even the uncertainty treatment [17]. Then, approaching the CDF around the wind power realization q through the first term of Taylor's series, it is obtained that:

$$F(w) = F(q) + f(q)(w - q) \tag{8}$$

Solving (7) and by replacing (8), it is obtained that:

$$UIC(w_F) = (k_r + k_p)[F(w_F) + f(w_F)(w - w_F)] - K \tag{9}$$

where $K = k_r F(w_F^-) + k_p F(w_F^+)$. Then, the wind power uncertainty cost (UC) is defined here as the expected balancing cost due to the uncertainties in the realization w_F between w_F^- and w_F^+. From the fundamental theorem of calculus, the wind power uncertainty cost can be calculated as a function of the scheduled wind power P_w as [41]:

$$UC(P_w) = E[C(\Delta w_F)] = \int_0^{P_w} UIC(w_F) dw + \gamma \tag{10}$$

where γ is a constant. Developing (10) it is obtained that:

$$UC(P_w) = (k_r + k_p)\int_0^{P_w}[F(w_F) + f(w_F)(w - w_F)]dw - \int_0^{P_w} K dw + \gamma \tag{11}$$

To solve (11), notice that the only integration variable is ω while all other parameters are constants. In fact, $F(\omega_F)$ and $f(\omega_F)$ represent the CDF, and the PDF evaluated at the wind power realization ω_F, and K and γ are known constants. Therefore, the integration of Eq. (11) yields:

$$UC(P_\omega) = (k_r + k_p)\left[\frac{1}{2}f(\omega_F)\omega^2\Big|_0^{P_\omega} + F(\omega_F)\omega\Big|_0^{P_\omega} - f(\omega_F)\omega_F\omega\Big|_0^{P_\omega}\right] - K\omega\Big|_0^{P_\omega} + \gamma \tag{12}$$

After that, it is possible to write the wind power uncertainty cost as a quadratic function depending on the scheduled wind power like:

$$UC(P_w) = (k_r + k_p)\left\{\frac{1}{2}f(w_F)P_w^2 + [F(w_F) - f(w_F)w_F]P_w\right\} - KP_w + \gamma \tag{13}$$

notice that γ can be easily adjusted to make $UC(P_w) = 0$ when the imbalance risk is minimum. Additionally, when it is considered that $k_r = k_p = \lambda_{UC}$, i.e., when upward and downward reserves are equally penalized, it can be obtained from Eq. (13) that:

$$UC(P_w) = \lambda_{UC}\left\{\frac{1}{2}f(w_F)P_w^2 + [F(w_F) - F(w_F^-) - F(w_F^+) - f(w_F)w_F]P_w\right\} + \gamma \tag{14}$$

In any case, the uncertainty cost function for each realization may be written as a quadratic function like:

$$UC(P_w) = \alpha P_w^2 + \beta P_w + \gamma \tag{15}$$

where α and β can be obtained comparing the Eqs. (14) and (15):

$$\alpha = \frac{1}{2}\lambda_{UC} f(w_F) \tag{16}$$

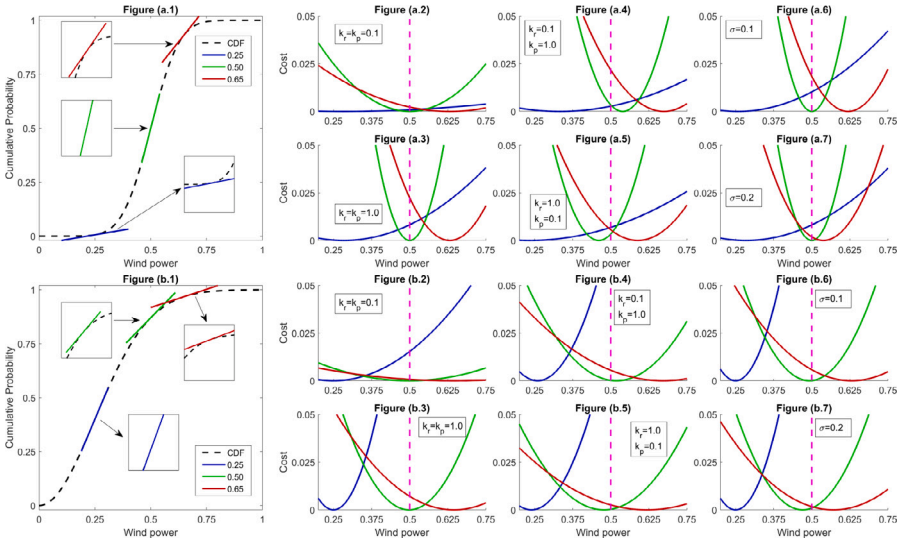

Fig. 2. Uncertainty cost function comparison for Normal and Weibull CDF.

$$\beta = \lambda_{UC}[F(w_F) - F(w_F^-) - F(w_F^+) - f(w_F)w_F] \tag{17}$$

As it can be seen from Eqs. (17) and (18), the values of α and β can be computed by evaluating the CDF and the PDF in the forecasted value ω_F and the extremes of the deviation intervals ω_F^- and ω_F^+ which are all known parameters. In contrast, γ could be adjusted to make the quadratic cost function equal to zero when the imbalance risk is minimal.

To summarize, this paper presents a generalization of the definition of the uncertainty cost proposed in [8]. While in [8] the entire CDF is approached through a piece-wise function of three segments, our model enables the estimation of the imbalance cost for any interval and around each possible realization on the wind power CDF. This generalization allows for better capture of the features and asymmetries of the imbalance risk at various distribution intervals, making it highly suitable for managing the high uncertainty of wind power in the medium- and long-term. Moreover, the mathematical formulation to properly estimate the uncertain cost function in this context is one contribution of this paper. This formulation ensures the proper quantification and representation of the imbalance cost considering intervals for the uncertainties associated with wind power generation and the reserves-cost management.

As an illustrative example, Fig. 2 shows the effect of approaching a Normal (mean 0.5 and standard deviation 0.1) and a Weibull CDF (scale parameter 0.35 and shape parameter 2.0) in three different realizations. Notice that the wind power has been normalized in magnitude. The approach has been evaluated assuming realizations of 0.25 (blue line), 0.50 (green line), and 0.65 (red line) for each CDF. Figure (a.1) and (b.1) shows the linear approaches made on Normal and Weibull CDF. Figures (a.2) to (a.7) and (b.2) to (b.7) illustrate the uncertainty cost functions for Normal and Weibull distributions, respectively, under different scenarios. For Figures (a.2) to (b.5), $\sigma = 0.05$, while for Figures (a.6) to (b.7), $k_r = k_p = 1.0$. Additional details are provided within each figure for further reference and analysis.

It can be seen from Fig. 2 that:

1. The minimum of the quadratic function is always close to the realization value, indicating that our approach effectively captures the proximity of deviations to the actual realization.
2. Deviations are more heavily penalized when they occur within the quasi-linear interval of the CDF. It can be seen from the tighter curves observed for realizations falling within these sections (green and blue lines in the Normal and Weibull CDFs, respectively). This behavior is desirable due to the higher probability of occurrences in these parts of the CDFs. In other words, when the imbalance risk is low, a higher penalty is imposed on values deviating from the forecasted ones. Also, notice that the shape of the curve closely relates to the parameter α, with higher values of α resulting in increased penalization.
3. As it was expected, Figures (a.2) to (b.3) demonstrate that higher balancing reserve prices lead to greater penalties for deviations.
4. Figures (a.4) to (b.5) show that if the upward reserve cost is less than the downward, the minimum of the curve shifts towards the right. This indicates that scheduled wind power tends to be higher to mitigate the risk of incurring additional costs associated with utilizing downward reserves. A similar analysis can be conducted if the cost of upward reserves exceeds that of downward reserves.
5. It can be seen from Figures (a.6) to (b.7) how the selection of the deviation interval can impact both the location of the minimum and the level of penalization, being less susceptible to realizations on the quasi-linear part of the CDF.

In a nutshell, the uncertainty cost function allows modeling the short-term balancing actions that system operators could take in the short term while modeling properly the imbalance risk faced in each realization. If a realization has a high probability of occurrence, the variations of this value are more penalized. It is a matter of fact that deviations from this realization would more likely produce imbalances. In contrast, those variations are more flexible for realizations with a low probability of occurrence. In this way, our approach differentiates the possible imbalance risk of the realizations and could be easily included in medium-term generation planning models. In addition, this approach can provide more flexibility to optimization techniques in those realizations with a low probability of occurrence, enabling the technique to find cheaper solutions than those obtained with fixed values approaches.

2.4. Temporal aggregation

Reducing the temporal scope of the problem is a commonly employed strategy for simplifying the complexity in medium- and long-term planning models. This practice is valuable when dealing with large-scale problems where many scenarios are needed to model the uncertainty properly, making it challenging to handle variables with a high level of granularity. A detailed review of different temporal aggregation methods applied to power systems analysis can be found

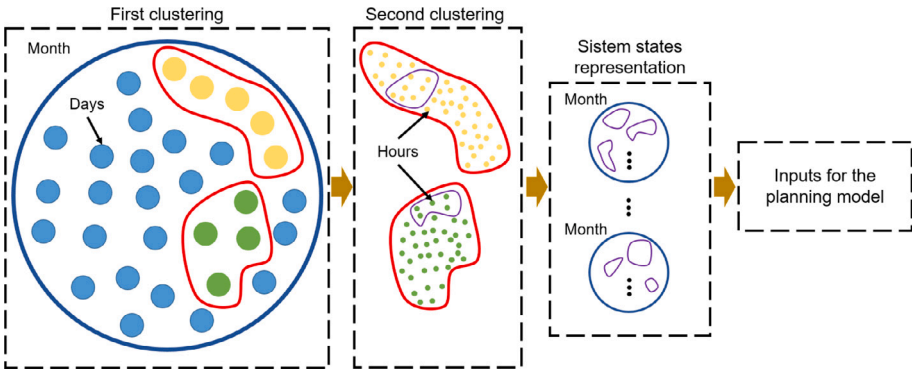

Fig. 3. Schematic representation of the proposed temporal aggregation.

in [42]. This paper has conducted a two-steps temporal aggregation approach inspired in [30,43]. First, the demand covered by dispatchable generation has been selected as the control variable for grouping similar operational hours. This selection allows capturing the spatial and temporal dynamics of the system while considering the behavior of the resources that must be scheduled on a mandatory basis, like, for instance, renewable energy sources, units that should be dispatched with a minimum production for reliability purposes, and nuclear generators. For this, initially, a k-means clustering algorithm has been performed to group the days of each month based on how similar they are and to select a reduced number of types of day representatives of the month [44]. A monthly basis has been chosen here because of the time resolution of different variables of midterm planning models like hydro management and fuel contracts. However, other temporal aggregations are admitted by the proposed methodology.

After that, a second clustering process was carried out to group the similar hours of the representative days into system states. With this approach, each state usually represents more than one hour, and each type of day is treated similarly. Thus, the computational burden of the medium-term planning model is reduced while preserving the market dynamics. In the case of the uncertainty cost functions, one quadratic function is calculated first for each hour. After that, the mean of the quadratic functions belonging to the grouped hours in the state is calculated, i.e., one quadratic function is computed for each state. This approach allows preserving, at the minimum of the function, the average of the forecasted values of the clustering hours while better capturing the average features of the quadratic functions, i.e., symmetry and convexity.

Temporal aggregation approaches have inherent limitations like they could lead to the loss of crucial temporal details which may include critical information, mainly when dealing with events that occur on a short time scale, like sudden storms and power outages. Nevertheless, this limitation can be addressed in our proposed methodology by defining a unique state for each hour. This hourly granularity proves a precise and comprehensive representation of imbalance risk and unexpected events. In this context, the trade-offs between computational efficiency and result accuracy should be carefully considered when defining the temporal aggregation's granularity for a particular study case.

Finally, once the uncertainty cost functions are calculated, they are adequately scaled according to the installed wind power capacity. The proposed temporal aggregation is schematically shown in Fig. 3.

2.5. Medium-term generation planning model

The proposed methodology has been implemented as an extension of the market equilibrium model developed in [45,46]. A perfect competition market has been simulated. The quadratic cost functions related to the uncertainty cost and the temporal aggregation have been included in the formulation. In this way, the medium-term generation planning can be determined by solving the following quadratic-optimization problem:

$$\min_{P_{fms}} \sum_{f=1}^{F}\sum_{m=1}^{M}\sum_{s=1}^{S}\left(C_{fms}(P_{fms}) + \frac{\theta_{fms}}{2}P_{fms}^{2} + UC(P_{fms})\right) \tag{18}$$

subject to:

$$\sum_{f}^{F} P_{fms} = D_{ms} \quad \forall m, s \tag{19}$$

$$P_{fms}^{F}(1-\sigma) \le P_{fms} \le P_{fms}^{F}(1+\sigma) \quad \forall f, m, s \tag{20}$$

$$\mathcal{H}(P_{fms}) \ge 0 \quad \forall f, m, s \tag{21}$$

where f, m, and s denotes firms, months, and states. Likewise, F, M, and S indicate the total number of firms, months, and system states. Besides, P_{fms}, P_{fms}^{F} and C_{fms} are the generated power, the forecasted wind power, and the production cost of the firm f, at the month m, and the state s. The production cost function is the sum of the variable thermal production costs, these units' start-up and shut-down costs, and the fixed cost of dispatched units, following the methodology proposed in [44]. Additionally, θ_{fms} is the conjectured-price response which measures the change in the market price $\lambda_{m,s}$ concerning the changes in the production of each firm. That is:

$$\theta_{fms} = -\frac{\partial \lambda_{m,s}}{\partial P_{f,m,s}} \quad \forall f, m, s \tag{22}$$

Eq. (19) corresponds to the power balance constraint, and Eq. (20) is the constraint of the maximum deviation allowed concerning the wind power realization. Finally, Eq. (21) represents other technical and economic constraints considered in the model. To give some examples, the model includes restrictions related to the allocation of hydrothermal resources throughout the whole medium-term horizon, emissions allowances for thermal units, fuel contracts, and energy flow limits among different operative areas [45,46].

Finally, notice that committing the generation planning, including the uncertainty cost function, has advantages concerning committing fixed forecasted values for wind resources. First, wind flexibility can achieve cheaper operation points for the whole system. Second, the imbalance cost for deviations from the forecasted value to the committed value is covered. If the actual wind power is the expected value, there will be no imbalance cost. Moreover, third, the model covers the deviation costs on both sides at the same time. For instance, if the commitment is under the forecasted value and the actual realization is above it, a part, if not all, of the deviation cost is covered and considered in the model output.

3. Results

The proposed methodology has been tested in a real-size study case. For this, the power markets of Spain, Portugal, and France have been simulated from 1 January 2022 to 31 December 2024, following the methodology described in Section 2.1. There are notable characteristics that make this study case interesting to be analyzed. On the one hand, each country has a unique combination of electrical sources. The total generation capacity installed for Portugal, Spain, and France is currently close to 19 GW, 112 GW, and 141 GW, respectively. From these values, 28%, 24%, and 14%, respectively, came from wind resources in 2022 [47]. Due to these differences, it is possible to capture a wide range of uncertainties from a local and global perspective. On the other hand, there are physical interconnections between Spain and France and between Spain and Portugal, with no physical interconnection between Portugal and France. Spain and Portugal's electricity markets are fused in the Iberian Electricity Market (MIBEL), and energy transfers between MIBEL and France depend on each market's rules. In this way, it is interesting to simulate the response of each market and the whole under different reserve-cost scenarios.

This section is ordered as follows. In Section 3.1, the aspects related to the treatment of the wind power uncertainty are discussed, i.e., the steps from Sections 2.1 to 2.4 of the methodology are detailed. For the sake of clarity, the results of the temporal aggregation are introduced before the uncertainty cost function calculations. Likewise, Section 3.2 shows the impact of the methods proposed to treat the wind power uncertainty on the generation-planning model (Section 2.5 of the methodology) considering deterministic and Monte Carlo simulations of three hundred scenarios.

3.1. Treatment of the wind power uncertainty

3.1.1. Scenarios generation

Hourly paths of wind power have been generated following the methodology proposed in [31]. Real wind power data from 2008 to 2020 for Spain, and 2012 to 2020 for Portugal and France, have been firstly treated. Subsequently, three hundred scenarios of wind power generation for each hour and country have been simulated from 2022 to 2024. Fig. 4 shows the CDF and the box plot representing the per-unit wind power obtained for each country for all the simulated hours. Red, green, and blue lines show the behavior of Spain, Portugal, and France, respectively. The CDF shows similar behavior for wind power in Spain and France. However, Spain exhibits a higher utilization of the installed capacity in almost all the distribution intervals. In contrast, a higher imbalance risk for Portugal is expected from the low slope of the CDF and the heavy right tail. These observations are more evident in the accompanying box plot, which reveals similar distributions among the regions for values below the median, but notable asymmetries in Portugal due to its higher skewness. It is worth mentioning here that the accuracy of the forecasted scenarios is crucial in guarantying the reliability of the results obtained through the proposed methodology.

3.1.2. Distribution fitting

After generating the wind power scenarios, a distribution has been fitted for each hour following the methodology indicated in Section 2.2. Table 1 presents the number of hours fitted to a particular distribution by area. Specifically, 26.304 h corresponding to the simulated period (1 January 2022 to 31 December 2024) have been fitted for each area. Results show that logistic regression provides a better fit for most hours in Spain and France, also suggesting the presence of heavy tails in the distributions. Moreover, the kernel approach shows better fitting results for Portugal, as seen in the last column of the table.

In addition, MAE and MAPE evaluation metrics have been computed for each possible realization. To do this, we compared the fitted CDF with the empirical cumulative distribution function (ECDF), as described in [35], derived from each realization for each hour of the generated scenarios. Subsequently, we calculated the average MAE and MAPE values over each hour of the day for the forecasting period. Table 2 shows the values obtained for MAE and RMSE in each hour and area. Results show that the fitting process approximates the CDF of the scenarios generated in all the hours and areas properly, regardless of the hour of the day. In addition, the last row of the table shows the mean value, indicating the highest performance for Portugal and the lowest for Spain.

Table 1
Distribution fitting obtained by country.

	Normal	Weibull	Gamma	Beta	Logistic	Kernel
Spain	894	1752	650	85	21 922	1001
Portugal	375	1421	3507	789	2172	18 040
France	419	7179	1106	243	11 346	6011

Table 2
Values obtained for MAE and RMSE by hour and country.

Hour	MAE [%]			RMSE [%]		
	Spain	Portugal	France	Spain	Portugal	France
1	2.73	1.89	2.44	3.16	2.36	2.87
2	2.74	1.87	2.43	3.17	2.34	2.87
3	2.72	1.87	2.43	3.15	2.33	2.87
4	2.73	1.86	2.45	3.15	2.33	2.88
5	2.72	1.87	2.44	3.14	2.34	2.87
6	2.71	1.89	2.44	3.14	2.36	2.88
7	2.73	1.86	2.43	3.15	2.33	2.87
8	2.72	1.90	2.45	3.14	2.37	2.89
9	2.69	1.91	2.45	3.11	2.38	2.89
10	2.68	1.89	2.48	3.10	2.36	2.91
11	2.68	1.87	2.48	3.10	2.34	2.91
12	2.71	1.90	2.43	3.13	2.37	2.87
13	2.68	1.87	2.48	3.11	2.33	2.92
14	2.67	1.90	2.46	3.10	2.36	2.90
15	2.69	1.89	2.46	3.12	2.35	2.89
16	2.69	1.90	2.45	3.11	2.37	2.88
17	2.69	1.89	2.44	3.11	2.36	2.87
18	2.69	1.88	2.45	3.11	2.34	2.89
19	2.69	1.89	2.44	3.11	2.36	2.88
20	2.68	1.89	2.40	3.10	2.36	2.84
21	2.71	1.91	2.45	3.13	2.38	2.89
22	2.70	1.91	2.43	3.12	2.38	2.87
23	2.72	1.89	2.45	3.14	2.37	2.89
24	2.74	1.91	2.44	3.17	2.38	2.87
Mean	**2.70**	**1.89**	**2.45**	**3.13**	**2.36**	**2.88**

3.1.3. Temporal aggregation

Temporal aggregation helps to reduce the execution times and the computational complexity in large planning models. In this way, selecting the type of days and system states for each realization is crucial. In general, the lower the number of system states, the lower the complexity of the problem. However, the cost of this reduction is losing accuracy in results. Moreover, the strategy adopted to cluster similar operational hours may significantly impact the obtained results. The main objective of the temporal aggregation followed in this paper is to propose an approach to extend the proposed methodology for including wind power uncertainty in the generation planning models used to tackle real-size problems. However, identifying the optimal temporal aggregation is out of the scope of this paper, and it is suggested to be investigated in future works. In this context, five types of days representatives of each month have been selected to reduce the problem's size in this paper. The clustering has been performed according to the net demand (the part of the demand covered by dispatchable generation). Additionally, six system states by type of day have been built to represent the electricity market. These values of the kind of days and the system states have been shown to have a good equilibrium among computational complexity, the objective function's accuracy, and the main variables of planning models similar to the proposed one in this paper [30].

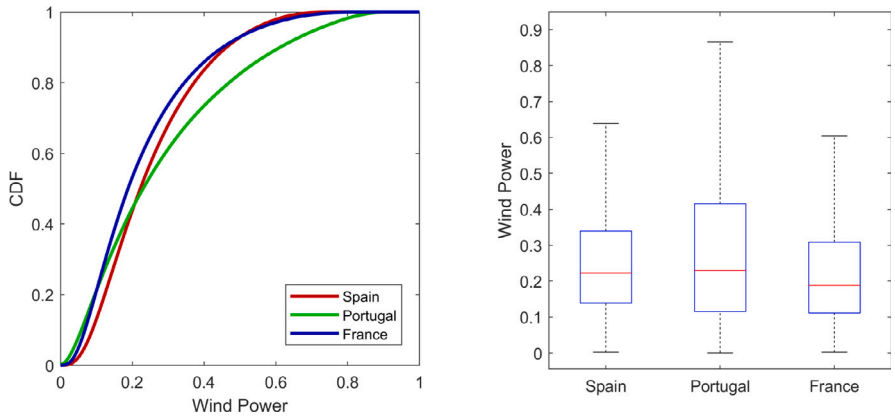

Fig. 4. CDF and box plot of wind power scenarios generated for Spain, Portugal, and France.

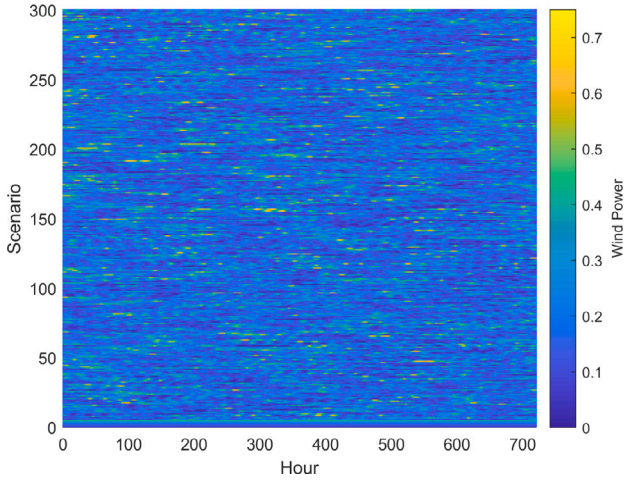

Fig. 5. Scenarios of wind power for June.

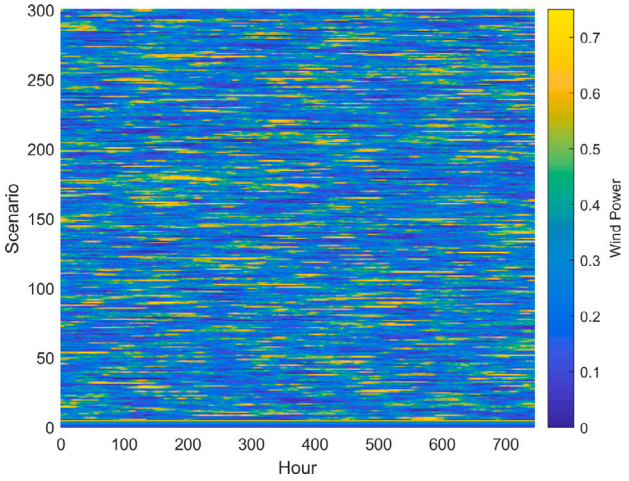

Fig. 6. Scenarios of wind power for December.

To illustrate the temporal aggregation performed, Figs. 5 to 10 show the wind power scenarios generated for June and December 2022 in Spain and their temporal aggregation. These months have been selected to show the different behavior in summer and winter. The horizontal axis indicates the hour of the month (for Figs. 5 to 8) or system states (for Figs. 9 and 10), depending on the figure. In contrast, the vertical axis shows the number of simulated scenarios (three hundred). One scale of colors on the right of each figure is provided to indicate the per-unit value of wind power in each realization or the number of the system states. Figs. 5 and 6 show the wind power scenarios generated for June and December. Notice that the wind power scale is the same for both Figures. It can be seen from those Figures a higher utilization in December from the more yellow and blue light colors in Fig. 6. In contrast, Figs. 7 and 8 show the belonging of the hours to each system state after the two-step temporal aggregation processes. From the dark-green columns of the figures, it can be seen that the hours have been grouped following a weekly pattern, approximately. It was expected because of the weekly seasonality of the demand captured by the k-means clustering algorithm. Finally, Figs. 9 and 10 indicate the wind power scenarios as a function of the system state representation.

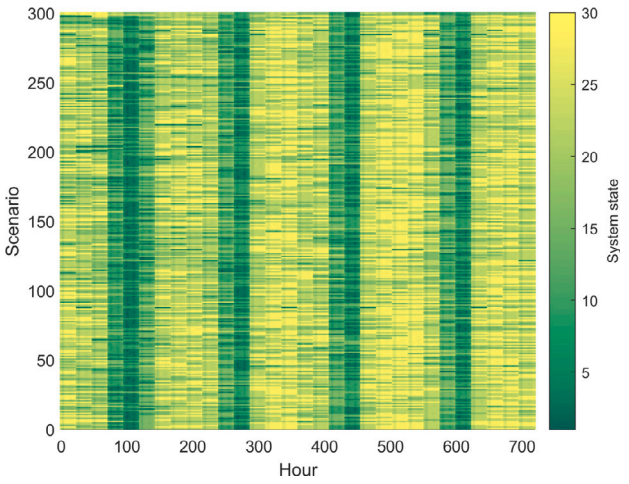

Fig. 7. Belonging of each hour of June to the system state.

3.1.4. Uncertainty cost

Once the distribution fitting is performed, the coefficients α, β, and γ from Eq. (15) are calculated for each hour. In this way, the uncertainty cost will be a function of the reserve costs and the deviation interval σ of each hour. In addition, one quadratic function is calculated to represent each system state properly and according to the temporal aggregation. To illustrate this, Figs. 11 and 12 show the quadratic cost functions obtained for Spain when $k_r = k_p = 20$, and $\sigma = 0.1$ for two of

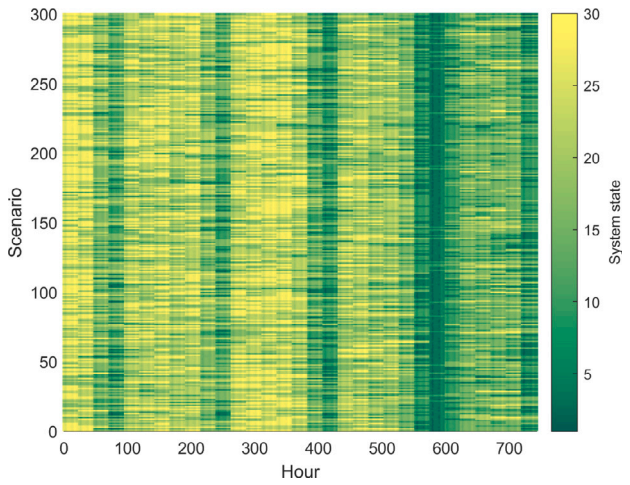

Fig. 8. Belonging of each hour of December to the system state.

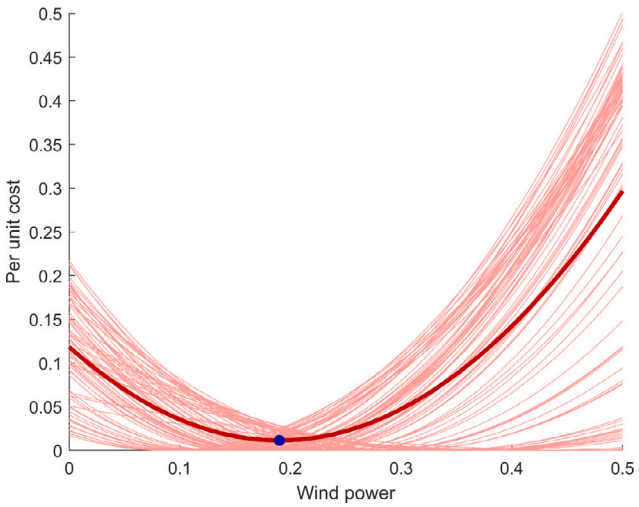

Fig. 11. Figure from June.

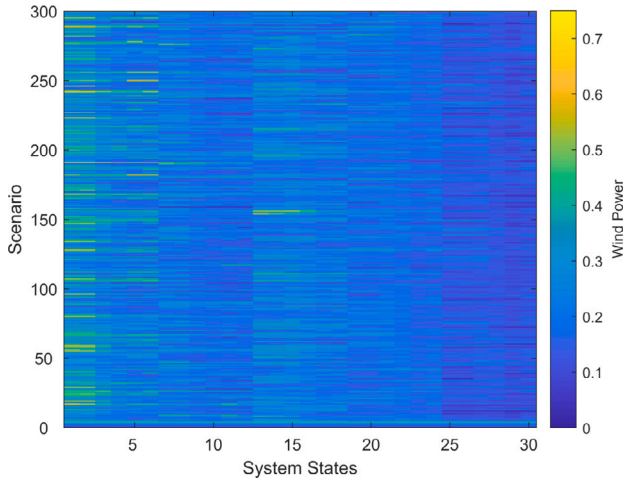

Fig. 9. Scenarios of wind power as a function of the system state representation for June.

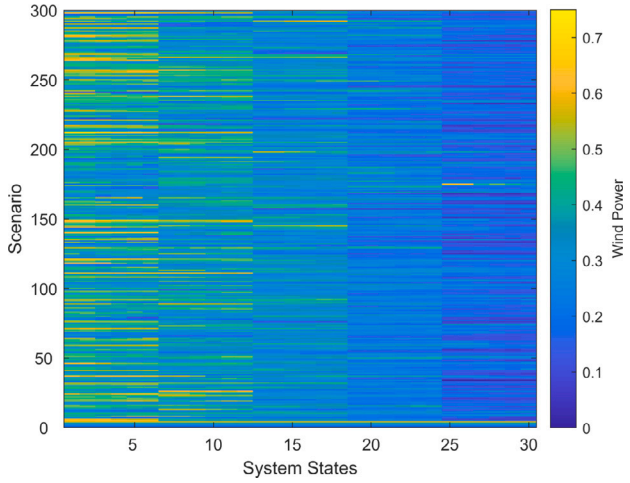

Fig. 10. Scenarios of wind power as a function of the system state representation for December.

the system states[1]. In particular, the thicker line on the figures shows the quadratic uncertainty cost obtained for the state with our approach. In particular, the hour 15:00 on the 21st of June and 21st of December 2022 has been evaluated. The hours mentioned pertain to system states 22 and 21, where a total of 81 h were aggregated in June and 72 h in December, respectively. The minimum values, represented by blue dots, are 0.19 and 0.25. These values indicate a higher value of wind energy expected for December. As it can be seen from the figures, the temporal aggregation proposed allows capturing the average minimum, symmetry, and convexity of the clustered quadratic cost functions for both months.

A close inspection of the curves indicates a higher increase rate for June, suggesting a lowest imbalance risk. It can also be seen from the quadratic coefficients α obtained for both curves, which were 4.0 for June and 2.3 for December, respectively. A fixed cost can be observed even if the actual wind power value matches the forecast. As it was suggested in Section 2.4, the value of γ can be adjusted to make the uncertainty cost equal to zero when there is no imbalance. Finally, Table 3 shows some descriptive statistics that summarize the values of the coefficients obtained for each area. A total of 360 uncertainty cost functions have been computed by area, being consistent to the five types of days of each month, and the six system states detailed in the temporal aggregation. The units of α, β, and γ are €/unit^2, €/unit, and €, respectively. As was expected from Fig. 4, α values suggest a higher imbalance risk for Portugal. Likewise, the highest deviation of all the constants is expected for France.

Finally, notice that each cost function properly represents the main characteristics of the state, as can be seen from Figs. 11 and 12. However, it is worth mentioning that when the quadratic functions are averaged to obtain one function for each state, some features of the uncertainty of each realization could get lost. This limitation can be addressed through our methodology by defining a higher number of states for each hour and, in a more accurate scenario, by defining a unique state for each hour.

[1] The values of k_r and k_p have been selected to be approximately realistic. For this, we subtract the time-average price of activated balancing energy mechanisms in the European Union electricity markets [48] with a typical offer price of a marginal conventional unit of the case of study. Likewise, allowed deviations have been chosen considering that they could change according to the market. A clear example is the PJM which establishes deviations exempts of payments from day-ahead energy bids of 5%. However, other countries can allow higher or lower deviations [40]. In this context, the deviation interval σ has been established in 0.1 to cover various scenarios exempt from payments and balancing capabilities.

Table 3
Descriptive statistics of the coefficients obtained by country.

Area/Coeff.	Mean			Std. deviation			Skewness			Kurtosis		
	α	β	γ	α	β	γ	α	β	γ	α	β	γ
Spain	2.31	−1.01	0.14	1.17	0.54	0.10	0.23	0.06	0.44	2.64	2.26	2.34
Portugal	1.67	−0.70	0.11	0.84	0.33	0.10	0.38	0.24	0.88	2.38	2.62	3.20
France	2.42	−0.87	0.11	1.26	0.43	0.08	0.26	0.02	0.74	2.59	4.10	3.23

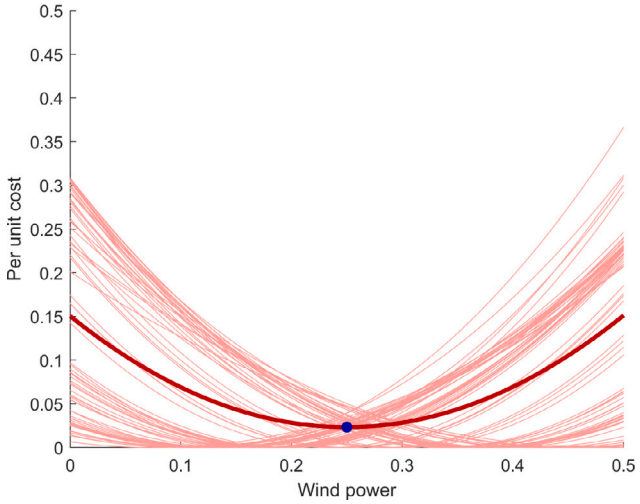

Fig. 12. Figure from December.

3.2. Impacts of the methodology on the medium-term generation planning model

This section has conducted simulations for the integrated markets of Portugal, Spain, and France from 1st January 2022 to 31st December 2024. Scenario- and sampling-based approaches have been used to evaluate the impact of including the imbalance risk in the medium-term generation planning model. In both approaches, we compare the market clearance when fixed forecasted realizations of wind power are considered (the traditional approach) and when the imbalance cost function is included in the market simulation (our proposal). Specifically, the variations of six parameters of the market clearance are analyzed: wind power production, average price, system operation cost, emissions cost, start and stop cost, and imbalance cost.

3.2.1. Case study 1: Scenario-based approach

This section compares the market clearance between the traditional approach and our proposal for percentiles P15, P30, and P50 of wind power generation for each hour. These specific percentiles were chosen as they represent scenarios that enable evaluating the CDF at different distribution intervals on the most linear part of the CDFs across all the areas (see Fig. 4). By selecting these percentiles, we aim to reduce the impact of significant asymmetries that may occur above the median in Portugal, simplifying the analysis of the results for this case study. That is because deviations from the forecasted values are expected to be lower within the quasi-linear interval of the CDF and, therefore, more comparable with the traditional approach. In contrast to wind power generation, all the other variables that could be subject to uncertainty (e.g., solar generation and demand) are fixed. The scenario-based approach allows evaluating the impact of our methodology in a simple way. Actually, it lets us isolate the effects of our methodology from other uncertain variables but wind power on the results. Finally, two scenarios have been proposed to be discussed. In the first scenario, the same reserve cost for all the countries is assumed; in the second one, the impact of having different reserve prices is analyzed. Those scenarios enable us to determine the effect of applying independent or joint prices for balancing mechanisms on interconnected electricity markets.

Scenario 1: Equal reserves' cost for all the areas

For this scenario, it has been assumed that $k_r = k_p = 20$ €/MWh and $\sigma = 0.1$ for all the countries. Table 4 shows the variations obtained from the market simulation when the generation planning model includes the imbalance risk. Variations have been denoted by the symbol Δ in the table. Changes in wind production, average price, costs of operation, emissions, start and stop, and imbalance are shown by country and percentile. In addition, the last row shows the sum of all the cost variations for easy reference. In particular, imbalance costs have been calculated considering the wind power deviations concerning the forecasted value. The units of the values are given in the first column. Positive values in production indicate a higher wind power generation obtained with our model. On the contrary, high positive values in the remaining entries indicate lower values from our approach. Then, it can see that our proposal increases the wind power production in all the percentiles. In contrast, it has reduced average prices for all the areas and percentiles. Likewise, our approach reduces the cost in all the considered percentiles but starts and stops costs in the P15 of Portugal and France, where they increase. However, this increase is minimal compared with earns achieved by reducing the other charges.

Furthermore, it can be seen that Spain and France show similar behavior for all entries in the table concerning the percentiles. Both improve differences as the percentile increases concerning the traditional approach, all variables but emissions cost, which has the highest reductions in the percentile P30. In contrast, all the entries achieve higher differences in percentile P30 for Portugal but imbalance cost, which has the highest reductions in the percentile P50. Notice that, contrary to what was expected, the differences in wind power productions in Portugal decay at the last percentile. In other words, our approach not always increases wind production with percentile to reduce the total cost of the planning. In this case, the reduction is given because the growth of wind power in Spain decreases its price and increases the energy flow from Spain to Portugal by 315 GWh. From the model perspective, importing this energy is cheaper than incurring an imbalance cost for increasing wind energy in Portugal.

Observe that the imbalance cost increases considerably with percentile, even when the differences with the wind power production are low (like P50 in Portugal). That is because of two reasons. First, the values shown for production in Table 4 are the sum of values for the entire planning horizon, which can be positive or negative depending on the hour. In contrast, the imbalance cost increases for both positive and negative deviations. Second, as it was shown in Eq. (20), the maximum power deviation allowed increases with a greater wind energy availability. In this way, results indicate significant positive and negative changes in hourly-wind energy production returning high imbalance costs in realizations with low probability of occurrence (like those found in percentile P50 for all the areas). Actually, results show that our approach captures and appropriately incorporates the imbalance risk of the realizations in the planning model by making wind power production more flexible in those unlikely realizations. This flexibility reduces the total cost in all the percentiles while incorporating the balancing costs in the medium-term simulations.

Scenario 2: Different reserves' cost for the areas

Notice that, in Scenario 1, the proposed uncertainty cost function considers the imbalance risk but does not account for the varying

Table 4
Variations when imbalance risk is included in simulations and $k_r = k_p = 20$ €/MWh for all the countries.

Var./Perc.	Spain			Portugal			France		
	P15	P30	P50	P15	P30	P50	P15	P30	P50
ΔProduction [MWh]	257 651	364 460	492 567	31 919	54 376	26 289	154 277	233 739	334 215
ΔAv. price [€]	0.42	1.55	1.65	0.03	1.16	0.93	0.27	0.79	1.26
ΔOp. cost [k€]	553 368	643 149	674 035	95 833	205 465	136 907	80 187	353 231	363 377
ΔEmi. cost [k€]	146 250	182 874	140 900	31 509	59 469	30 590	29 943	108 518	88 815
ΔSt&St cost [k€]	308	2664	3750	−1097	653	16	−529	690	842
ΔImb. cost [k€]	542 523	1 700 894	4 071 072	299 058	1 053 902	2 403 032	418 567	1 214 868	2 837 021
ΔTotal cost [k€]	1 242 450	2 529 582	4 889 757	425 304	1 319 490	2 570 545	528 168	1 677 308	3 290 055

Table 5
Variations when imbalance risk is included in simulations and $k_r = k_p = 40$ €/MWh for Portugal, and $k_r = k_p = 20$ €/MWh for Spain and France.

Var./Perc.	Spain			Portugal			France		
	P15	P30	P50	P15	P30	P50	P15	P30	P50
ΔProduction [MWh]	257 651	364 460	492 567	27 243	4898	−80 271	154 277	233 739	334 215
ΔPrice [€]	0.41	1.36	1.37	0.07	1.05	0.75	0.27	0.74	1.16
ΔOp. cost [k€]	551 333	579 724	552 271	89 351	180 097	118 887	79 707	332 892	332 809
ΔEmi. cost [k€]	145 411	162 479	116 983	28 748	50 904	26 557	29 761	102 564	82 455
ΔSt&St cost [k€]	372	2085	2526	−1049	706	−154	−470	627	609
ΔImb. cost [k€]	542 523	1 700 894	4 071 072	579 131	1 901 127	4 332 080	418 567	1 214 868	2 837 021
ΔTotal cost [k€]	1 239 641	2 445 183	4 742 853	696 181	2 132 835	4 477 370	527 566	1 650 872	3 252 895

weight that the installed capacity of wind power may have on each country's energy mix. This means that deviations of similar magnitude are treated equally for all countries, regardless of their installed capacity and resource dependency. However, in practice, countries with higher dependency on wind energy may have a stronger need to minimize forecasting deviations in order to ensure the secure operation of their power systems. In this scenario we explore a more restrictive forecasting deviation scenario for Portugal, which has the higher dependency on wind energy compared to the other countries of the study case. The results obtained under this more restrictive scenario will be compared with those obtained in Scenario 1. For this, it has been assumed that $k_r = k_p = 40$ €/MWh for Portugal. In contrast, the reserve prices of Spain and France have remained equal to 20 €/MWh. Likewise, the deviation interval has been preserved in 0.1 to make results comparable with were obtained in Scenario 1.

The results of this scenario are shown in Table 5. Observe that this table has the same arrangement as Table 4. It can see from both tables that the variations in wind power production and the total imbalance cost of Spain and France remain constant. Moreover, similar behavior is identified for both countries when the two scenarios are compared but operation cost, where the highest deviation is achieved at percentile P30 instead of P50. In particular, observe that the reduction in operational and emission cost values between the two tables suggests a higher net generation from conventional plants in Scenario 2. It is, of course, because the lower values in Table 5 indicate nearer values to the traditional approach, which has less net wind energy and high conventional generation to supply the demand.

In contrast to Spain and France, wind power production and total imbalance costs have changed in Portugal. As was expected, productions are more similar to forecasted for percentiles P15 and P30 because of the increase in reserve price that increases the imbalance penalization. However, contrary to the expectations, the reserves price increase has grown the differences concerning predicted values of wind power in percentile P50. Moreover, the negative sign indicates that the net wind power production is less than forecasted. In this percentile, the observed behavior in scenario one has intensified here because the high reserve prices in Portugal increase the differences between the prices of the Spanish and Portuguese markets even more. In this case, Portugal's differences in imports increase by around three times from the scenario 1, which also explains the increase in conventional production in Spain and France. Furthermore, it can be seen from the last row that almost all the reductions in cost achieved with our approach in the first scenario have been reduced but in Portugal. These results show that unilateral increases in the reserve prices in one country could not decrease deviation and decrease planning costs but could also negatively impact other markets.

3.2.2. Case study 2: Sampling-based approach

This section analyzes the differences obtained for the market clearance when a Monte Carlo simulation of three hundred scenarios is performed using the medium-term generation planning model. All the simulations have been conducted in a computer with 128 GB of installed RAM and 48 logical processors with Intel(R) Xeon(R) Silver 4116 CPU @2.10 GHz. With this machine, the scenarios' simulations took 7 h and 13 min for the traditional approach. In contrast, finishing all the scenarios with our approach took 6 h and 37 min, i.e., 36 min of difference or saving of computational time close to 8%. These savings in computational time can be attributed to the flexible representation of wind power generation in our approach. By relaxing the constraints associated with the balance between generation and demand, our methodology alleviates the computational burden on the optimization problem. As a result, the optimization solver exhibits improved performance, leading to more efficient and faster market clearance.

The same six variables described in Section 3.2 are compared, with $\sigma = 0.1$ and $k_r = k_p = 20$ €/MWh in the three countries. In contrast to the first case study, our simulations in this study incorporate realistic wind power paths throughout the planning horizon, rather than relying on percentile scenarios. Furthermore, we have included realistic paths for demand and solar generation. By adopting a sampling-based approach, we can assess the impact of our proposed methodology on market clearance in a more uncertain environment, resembling real-world medium-term scenarios. This extends beyond the scope of the simulations conducted in the initial case study. Six CDFs have been fitted for each analyzed variable to show the variations regarding the traditional approach obtained in this case study. Results obtained for the CDF of each of the six cost variation analyzed in case study 1 are shown in Fig. 13. Likewise, the CDF obtained for the total cost variations are shown in Fig. 14 for an easy reference. In both figures, red, green, and blue lines indicate results obtained for Spain, Portugal, and France, respectively. Likewise, each variable and its units are indicated on the horizontal axis.

The analysis of these figures reveals some key findings about our approach. Firstly, it can be seen from the production distributions that our approach leads to an increase in wind power production in Spain and France, while its impact on production in Portugal exhibits some

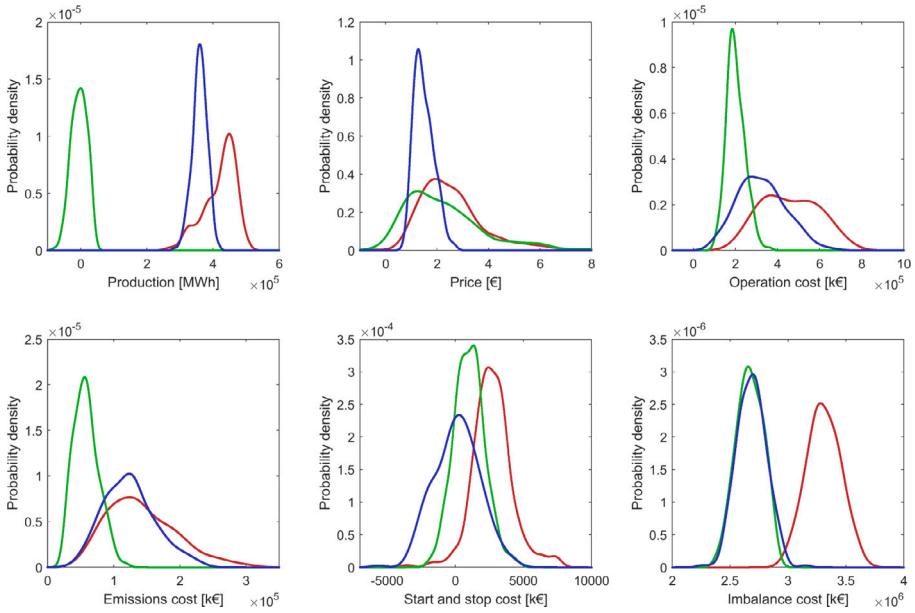

Fig. 13. PDF for the variables considered in Monte Carlo simulations.

reductions. It can be seen from the part of the CDF of the production curve that falls in negative variations. The realistic paths used to model the uncertainty of wind power, demand, and solar generation play a significant role in explaining the observed discrepancies in Portugal's production. The analysis of price distribution further supports this observation, highlighting a more significant decrease in Spanish electricity prices compared to Portugal. Consequently, this price difference fosters an augmented energy flow from Spain to Portugal, as previously discussed in the first case study.

Additionally, our methodology consistently achieves a reduction in operation and emissions costs, with results predominantly clustered around the mean in the case of Portugal. Remarkably, despite the decrease in wind power production, Portugal attains lower operating costs. This outcome can be attributed to the combination of reduced reliance on fossil technologies within Portugal and an increased import of energy from Spain, leading to more favorable operating cost outcomes. Furthermore, the analysis of start and stop costs reveals that our approach could reduce or increase these costs across all areas. However, it is notable that our methodology demonstrates a reduction for Spain in almost all the scenarios. In contrast, Spain has the higher imbalance cost. Remarkably, the imbalance cost in France is more similar to Portugal than Spain, despite the fact that the change in net wind power production in France is more similar to Spain. This discrepancy indicates that deviations from forecasted values are more frequent and significantly lower in France compared to Portugal. This can be attributed to the non-linear penalization introduced by the quadratic function, which leads to significantly different imbalance costs for the same net wind power deviations.

Lastly, Fig. 14 provides evidence that the total cost is reduced for all countries, including Portugal, despite experiencing a decrease in wind power in specific scenarios. This underscores the overall cost reduction benefits of our proposal across all the analyzed countries.

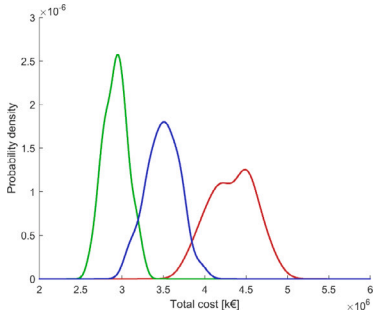

Fig. 14. PDF for the total cost variation obtained in Monte Carlo simulations.

4. Conclusions

This paper proposes a novel methodology for incorporating wind energy imbalance risk into medium-term generation planning models. By evaluating a realistic study case involving the interconnected electricity markets of Spain, Portugal, and France, we have demonstrated the effectiveness of our methodology. Our overall findings indicate that incorporating an imbalance risk that depends on wind power distributions enhances the flexibility of wind power production in instances with a low probability of occurrence, alleviates constraints related to the balance between generation and demand, and induces changes in various outcomes of the generation planning models. In addition, we show that the transmission grid notably influences these changes. The results indicate that our representation of the wind power uncertainty adequately approximates the forecasted values in all the hours and countries. Furthermore, the proposed temporal aggregation approach has successfully captured the average characteristics of the wind power uncertainty representation for different months, resulting in fewer scenarios for evaluation. Additionally, based on the outcomes of the study cases, our approach accurately models wind power imbalance risk at different distribution percentiles while simultaneously lowering planning costs. Moreover, it reduces computational times by 8% when applied to our realistic study case. Also, we show that a planning strategy oriented to reducing the imbalance risk could mean that a country should reduce its expectations of generating with renewable resources. Finally, our analysis has shed light on the impact of increasing the penalization price for imbalances in a country. Our analysis indicates that increasing the penalization price for imbalances in a country does not necessarily reduce the imbalances and decrease planning costs but could also negatively impact other markets. This finding underscores the importance of considering interconnections when defining reserve incentives and penalties for balancing in multi-area electricity markets. However, the applicability of our results may be influenced by the

unique geographical and climatic features of the case study. Future research efforts should broaden the scope by including data from more countries and regions, allowing for a more comprehensive assessment of the methodology's applicability and robustness. In addition, we encourage exploring and implementing alternative methods at each stage of our methodology to enhance our proposal further.

CRediT authorship contribution statement

Geovanny Marulanda: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Antonio Bello:** Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. **Javier Reneses:** Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

References

[1] Sadorsky P. Wind energy for sustainable development: Driving factors and future outlook. J Clean Prod 2021;289:125779.
[2] Fang X, Hodge B-M, Du E, Kang C, Li F. Introducing uncertainty components in locational marginal prices for pricing wind power and load uncertainties. IEEE Trans Power Syst 2019;34(3):2013–24.
[3] Agency IRE. Future of wind: deployment, investment, technology, grid integration and socio-economic aspects. Tech. rep., IRENA; 2019.
[4] Musial W, Spitsen P, Beiter P, Duffy P, Marquis M, Cooperman A, Hammond R, Shields M. Offshore wind market report: 2021 edition. tech. rep, us department of energy (doe), energy efficiency & renewable energy (eere), Office of Energy Efficiency & Renewable Energy; 2021.
[5] Dadashi M, Zare K, Seyedi H, Shafie-khah M. Coordination of wind power producers with an energy storage system for the optimal participation in wholesale electricity markets. Int J Electr Power Energy Syst 2022;136:107672.
[6] Li J, Zhou J, Chen B. Review of wind power scenario generation methods for optimal operation of renewable energy systems. Appl Energy 2020;280:115992.
[7] Sarathkumar TV, Banik A, Goswami AK, Dey S, Chatterjee A, Rakshit S, Basumatary S, Saloi J. Uncertainty borne balancing cost modeling for wind power forecasting. Energy Sources B 2019;14(7–9):291–303.
[8] Yan J, Li F, Liu Y, Gu C. Novel cost model for balancing wind power forecasting uncertainty. IEEE Trans Energy Convers 2016;32(1):318–29.
[9] Holttinen H. Wind integration: experience, issues, and challenges. In: Advances in energy systems: The large-scale renewable energy integration challenge. Wiley Online Library; 2019, p. 341–54.
[10] Li Y, Hu B, Niu T, Gao S, Yan J, Xie K, Ren Z. GMM-HMM-based medium- and long-term multi-wind farm correlated power output time series generation method. IEEE Access 2021;9:90255–67.
[11] Xiao D, Chen H, Wei C, Bai X. Statistical measure for risk-seeking stochastic wind power offering strategies in electricity markets. J Mod Power Syst Clean Energy 2021;10(5):1437–42.
[12] Xu B, Zhu F, Zhong P-a, Chen J, Liu W, Ma Y, Guo L, Deng X. Identifying long-term effects of using hydropower to complement wind power uncertainty through stochastic programming. Appl Energy 2019;253:113535.
[13] Zakaria A, Ismail FB, Lipu MH, Hannan MA. Uncertainty models for stochastic optimization in renewable energy applications. Renew Energy 2020;145:1543–71.
[14] Singh V, Moger T, Jena D. Uncertainty handling techniques in power systems: A critical review. Electr Power Syst Res 2022;203:107633.
[15] Jordehi AR. How to deal with uncertainties in electric power systems? A review. Renew Sustain Energy Rev 2018;96:145–55.
[16] Soroudi A, Amraee T. Decision making under uncertainty in energy systems: State of the art. Renew Sustain Energy Rev 2013;28:376–84.
[17] Aien M, Hajebrahimi A, Fotuhi-Firuzabad M. A comprehensive review on uncertainty modeling techniques in power system studies. Renew Sustain Energy Rev 2016;57:1077–89.
[18] Roald LA, Pozo D, Papavasiliou A, Molzahn DK, Kazempour J, Conejo A. Power systems optimization under uncertainty: A review of methods and applications. Electr Power Syst Res 2023;214:108725.
[19] Ahmed SD, Al-Ismail FS, Shafiullah M, Al-Sulaiman FA, El-Amin IM. Grid integration challenges of wind energy: A review. IEEE Access 2020;8:10857–78.
[20] Neniškis E, Galinis A. Representation of wind power generation in economic models for long-term energy planning. Energetika 2018;64(1).
[21] Khorramdel B, Zare A, Chung C, Gavriliadis P. A generic convex model for a chance-constrained look-ahead economic dispatch problem incorporating an efficient wind power distribution modeling. IEEE Trans Power Syst 2019;35(2):873–86.
[22] Berahmandpour H, Kouhsari SM, Rastegar H. A new flexibility based probabilistic economic load dispatch solution incorporating wind power. Int J Electr Power Energy Syst 2022;135:107546.
[23] Zhang Y, Gu C, Yan J, Li F. Cournot game based multi-supplier local energy trading. Energy Procedia 2019;158:3528–33.
[24] Zhang Y, Gu C, Yan X, Li F. Cournot oligopoly game-based local energy trading considering renewable energy uncertainty costs. Renew Energy 2020;159:1117–27.
[25] Xu H, Chang Y, Zhao Y, Wang F. A new multi-timescale optimal scheduling model considering wind power uncertainty and demand response. Int J Electr Power Energy Syst 2023;147:108832.
[26] Wang Q, Luo X, Ma H, Gong N. Multi-stage stochastic wind-thermal generation expansion planning with probabilistic reliability criteria. IET Gener Transm Distrib 2022;16(3):517–34.
[27] Morales-España G, Lorca Á, de Weerdt MM. Robust unit commitment with dispatchable wind power. Electr Power Syst Res 2018;155:58–66.
[28] Ning C, You F. Data-driven adaptive robust unit commitment under wind power uncertainty: A Bayesian nonparametric approach. IEEE Trans Power Syst 2019;34(3):2409–18.
[29] Li P, Yang M, Wu Q. Confidence interval based distributionally robust real-time economic dispatch approach considering wind power accommodation risk. IEEE Trans Sustain Energy 2020;12(1):58–69.
[30] Orgaz A, Bello A, Reneses J. Modeling storage systems in electricity markets with high shares of renewable generation: A daily clustering approach. Int J Electr Power Energy Syst 2022;137:107706.
[31] Marulanda G, Bello A, Cifuentes J, Reneses J. Wind power long-term scenario generation considering spatial-temporal dependencies in coupled electricity markets. Energies 2020;13(13):3427.
[32] Xie Z, Ji T, Li M, Wu Q. Quasi-Monte Carlo based probabilistic optimal power flow considering the correlation of wind speeds using copula function. IEEE Trans Power Syst 2017;33(2):2239–47.
[33] González-Longatt F, Wall P, Terzija V. Wake effect in wind farm performance: Steady-state and dynamic behavior. Renew Energy 2012;39(1):329–38.
[34] Massey Jr FJ. The Kolmogorov-Smirnov test for goodness of fit. J Amer Statist Assoc 1951;46(253):68–78.
[35] Jung C, Schindler D. Global comparison of the goodness-of-fit of wind speed distributions. Energy Convers Manage 2017;133:216–34.
[36] Vargas SA, Esteves GRT, Maçaira PM, Bastos BQ, Oliveira FLC, Souza RC. Wind power generation: A review and a research agenda. J Clean Prod 2019;218:850–70.
[37] Han Q, Ma S, Wang T, Chu F. Kernel density estimation model for wind speed probability distribution with applicability to wind energy assessment in China. Renew Sustain Energy Rev 2019;115:109387.
[38] Bozdogan H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika 1987;52(3):345–70.
[39] Willmott CJ. Some comments on the evaluation of model performance. Bull Am Meteorol Soc 1982;63(11):1309–13.
[40] Schneider I, Roozbehani M. Energy market design for renewable resources: Imbalance settlements and efficiency–robustness tradeoffs. IEEE Trans Power Syst 2017;33(4):3757–67.
[41] Rudin W. Real and complex analysis. New York, NY: McGraw-Hill Book Company; 1966.
[42] Sangadiev A, Gonzalez-Castellanos A, Pozo D. A review on recent advances in scenario aggregation methods for power system analysis. 2022, arXiv preprint arXiv:2207.09557.
[43] Orgaz A, Bello A, Reneses J. Temporal aggregation for large-scale multi-area power system models. IET Gener Transm Distrib 2022;16(6):1108–21.
[44] Wogrin S, Dueñas P, Delgadillo A, Reneses J. A new approach to model load levels in electric power systems with high renewable penetration. IEEE Trans Power Syst 2014;29(5):2210–8.
[45] Barquin J, Centeno E, Reneses J. Medium-term generation programming in competitive environments: a new optimization approach for market equilibrium computing. IEEE Proc-Gener Transm Distrib 2004;151(1):119–26.
[46] Reneses J, Centeno E, Barquin J. Coordination between medium-term generation planning and short-term operation in electricity markets. IEEE Trans Power Syst 2006;21(1):43–52.
[47] Entso-e. European network of transmission system operators. Transparency platform. 2023, https://transparency.entsoe.eu, accessed: 2023-06-28.
[48] Entso-e. ENTSO-E Balancing report 2022. 2023, accessed: 2023-06-28.