

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This paper appears in: IEEE Transactions on Sustainable Energy, Digital Object Identifier: 10.1109/TSTE.2023.3305865

Uncertainty-Aware Procurement of Flexibilities for Electrical Grid Operational Planning

Ricardo J. Bessa, *Senior Member, IEEE*, Ferinar Moaidi, João Viana, and José R. Andrade

Abstract—In the power system decarbonization roadmap, novel grid management tools and market mechanisms are fundamental to solving technical problems concerning renewable energy forecast uncertainty. This work proposes a predictive algorithm for procurement of grid flexibility by the system operator (SO), which combines the SO flexible assets with active and reactive power short-term flexibility markets. The goal is to reduce the cognitive load of the human operator when analyzing multiple flexibility options and trajectories for the forecasted load/RES and create a human-in-the-loop approach for balancing risk, stakes, and cost. This work also formulates the decision problem into several steps where the operator must decide to book flexibility now or wait for the next forecast update (*time-to-decide* method), considering that flexibility (availability) price may increase with a lower notification time. Numerical results obtained for a public MV grid (Oberrhein) show that the *time-to-decide* method improves up to 22% a performance indicator related to a cost-loss matrix, compared to the option of booking the flexibility now at a lower price and without waiting for a forecast update.

Index Terms—Uncertainty, forecasting, risk, electrical grid, flexibility, decision-making.

I. INTRODUCTION

IN recent years, the integration levels of renewable energy sources (RES) have been steadily increasing since the concern about global warming and energy dependency pushed nations worldwide to set ambitious RES targets. In this context, Holttinen et al. reported RES curtailment and congestion problems in several power systems [1], e.g., wind energy dispatch-down level in Ireland rising from 5.3% to 11.4%; 30% increase of RES curtailment in Chile due to grid congestion and inflexible thermal generators; significant transmission constraints remain in Germany due to delayed grid reinforcements. A survey from CIGRE showed that RES is expected to have more impact on the need for short-term flexibility (from 15-min to 12-hours) compared to real-time and very short-term time horizons [2].

The literature about modern grid planning methodologies and ongoing revisions of the regulatory frameworks consider the use of active measures (flexibility from local grid resources) to postpone traditional grid investments while accommodating ambitious RES targets at the same time [3].

Ricardo J. Bessa, João Viana, and José R. Andrade are with Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto, Portugal, emails: {ricardo.j.bessa; joao.p.viana, jose.r.andrade}@inesctec.pt. Ferinar Moaidi is with INESC TEC and Faculty of Engineering of the University of Porto (FEUP), e-mail: ferinar.moaidi@inesctec.pt. The research leading to this work is part of the Smart4RES project (European Union's Horizon 2020, No. 864337). The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

This paradigm will require the design of flexibility markets for grid-centric services like congestion management and voltage control, which range from voluntary short-term procurement up to mid-/long-term tenders [4], complemented with regulated (or non-market) mechanisms like non-firm connection agreements, dynamic grid tariffs or bilateral contracts.

Therefore, it is necessary to revisit the traditional power system operating processes and software in control rooms, including human-machine interaction [5]. In particular, the high RES integration levels require a risk-aware decision-making structure where the activation of flexibility can be planned to ensure sufficient capacity to handle the forecasted technical issues. In this sense, it is crucial to provide fast decision-aid to operators and reduce the volume of information, particularly under load and RES forecast uncertainty.

A. Literature Review

The optimal power flow (OPF) method (and its variants) is the standard approach for short-term management of grid flexibility. With the impact of RES forecast uncertainty in the grid operational planning, the risk-based OPF and OPF under forecast uncertainty problems are timely and relevant and, in [6], the state-of-the-art methods, are categorized as a) risk-neutral and risk-averse two-stage stochastic optimization, b) chance-constrained optimization, c) robust optimization, and d) distributionally robust optimization. For the sake of completeness, we present several OPF-based approaches below.

In stochastic and chance-constrained (CC) optimization, the uncertainty is represented via a probability distribution. Zhang and Li proposed a CC-OPF with a back-mapping approach and a linear approximation of the power flow equations where uncertainty follows a multivariate Gaussian [7]. Roald et al. described a CC-OPF model solved using a randomized optimization technique, representing RES and load forecast uncertainty via scenarios [8]. This work was generalized to the AC OPF problem in [9]. Mezghani et al. proposed a data-driven method based on sparse regression to reduce the number of scenarios required to solve the stochastic AC OPF [10].

Robust optimization does not make any specific assumptions on probability distributions, and the uncertain parameters are assumed to belong to a deterministic uncertainty set. Soares et al. described a two-stage robust AC OPF for a distribution system operator (DSO) contracting market-based DER flexibility [11], where uncertainty is defined as the convex hull. Guo et al. considered ambiguity sets for multi-period control policies robust to forecast and sampling errors, using different linearizations of the AC power flow equations [12].

Reinforcement learning is also emerging as an alternative to traditional mathematical optimization for the OPF

problem, when formulated as a sequential Markov decision process [13]. Yet, integration of forecast uncertainty remains an open challenge for RL, at least for large-scale grids.

The present paper follows a different direction from the OPF-based methods and, instead of formulating a mathematical optimization problem, it proposes a decision-making sequence where the human operator is directly involved in the interpretation of action-cause relations leveraging from sensitivity indices, and its attitude towards risk is integrated to find the preferred solution. This approach is aligned with the industry's need to have simple and easily interpretable control rules and a limited number of actions, such as the MV generation curtailment algorithm, developed by EDF R&D for Enedis (French DSO) to deliver day-ahead constraints forecast and generators' limitation calculation until the real-time set-points are sent to generators [14]. An approach based on sensitivity indices was proposed in [15] to forecast which RES power plants must be curtailed and to what extent and ranks the flexible resources solely based on the sensitivity value.

The present paper also formulates the decision-making problem in a *time-to-decide* fashion where the human operator must decide to i) book flexibility now (and pay an availability price) or ii) wait for the next forecast update and book flexibility later (if necessary). This decision problem has been around in other domains, and one illustrative example is the time-dependent version of the cost-loss ratio as described by Murphy and Ye in [16]. Other examples are: Wanke and Greenbaum, when exploring airspace congestion resolution, formulated a three-stage decision tree to find the optimal time and type of action [17]; Jewson et al., in meteorological forecasts, proposed an extended cost-loss model to decide whether to base a decision on the first forecast or to wait for the second forecast [18]. The same authors also suggested a method of effectively conveying information regarding forecast changes to human decision-makers [19]. This was achieved by introducing different metrics, e.g., mean absolute change and probability of change of size x .

In the energy domain, specifically in terms of energy trading, Tankov and Tinsi proposed stochastic differential equations to describe the evolution of the forecast error as new information becomes available [20]. Bellenbaum et al. designed a two-stage stochastic regional flexibility market mechanism that integrates the trade-off between early procurement at low-cost and later procurement at a higher cost but with better forecasts and knowledge about the flexible load [21]. Mühlpfordt et al. established a relation between the *price of uncertainty* and the total variational distance between the densities of the in-hindsight OPF and CC-OPF solutions [22]. The *price of uncertainty* can be used to decide between solving technical problems in real-time or in predictive mode.

B. Contribution

Compared to the state-of-the-art discussed above, the main contributions from this paper are the following:

- A predictive grid management framework that uses sensitivity indices and risk metrics to rank flexibility options and provide a summarized view to human operators

in multi-criteria problems, taking the form of risk vs cost curves. This approach stands in contrast to OPF-based methods [6] that, despite their mathematical interpretability, do not enable human operators to analyze action-cause relations extracted from sensitivity indices. Furthermore, it facilitates enhanced interaction between operators and the decision-aid tool, specifically by allowing them to rank and explore various flexibility solutions and adjust the risk level based on the decision stakes. This relationship between risk level and stakes has not been previously studied in prior works about stochastic [8], [9] and robust optimization [11], [12]. As opposed to [14], the proposed approach includes information about forecast uncertainty and, conversely to [14], [15], it proposes a multi-criteria decision problem, ranking the flexible resources according to different criteria (instead of selecting resources solely based in the highest sensitivity).

- Introduces a *time-to-decide* formulation for the decision-making problem, which was overlooked in the state-of-the-art formulations [9]–[12], [14], [15], and that holds particular relevance in dynamic environments, where frequent forecast updates are necessary for adapting to changing conditions. While [21] introduces a flexibility market mechanism, it overlooks the methodology employed by the system operator to estimate flexibility needs. Thus, the proposed approach is complementary to [21] as it addresses this crucial aspect. Moreover, unlike [22], the proposed approach allows the operator to interpret risk-cost relations and make decisions that rely not solely on a distance metric. Additionally, this work presents a novel concept known as *second-level forecasting*. This novel approach involves employing a *second-level* model to forecast the uncertainty (represented by a set of quantiles) of future (uncertainty) forecasts. It was determined through experimentation that an encoder-decoder deep learning model showed higher accuracy than other approaches. This approach contrasts with [18], which applies a parametric approach inadequate for modeling RES forecast uncertainty.

C. Structure

The rest of this paper is organized as follows: Section II describes the conceptual framework; Section III describes how knowledge regarding flexibility modeling is derived for utilization in the decision-aid phase of Section IV; Section V presents the numerical results for the Oberrhein MV grid; the conclusions and future work are discussed in Section VI.

II. CONCEPTUAL FRAMEWORK

A. Flexibility market operation

Based on the reviewed commercial flexibility market platforms [4] and the survey recently conducted by the Joint Research Centre (JRC) with companies like EPEX SPOT or Enedis [23], this work assumes that both long-term and short-term products will probably be requested in future flexibility markets, subject to the specific network needs in each case.

The long-term flexibility products are designed to defer network investment, while the short-term flexibility products target congestion management and the strengthening of network resilience. However, a significant challenge lies in determining the appropriate price caps (or annual flexibility budget), which is comparatively more straightforward regarding long-term products associated with investment deferral [24].

According to JRC, while the present state of local flexibility markets in Europe does not provide a conclusive outlook regarding their future characteristics, a notable finding from the interviews is the anticipation of transitioning towards short-term flexibility markets [23]. The main reasons are [23]: a) by enabling smaller assets (e.g., EVs) to participate in the procurement process, there is enhanced liquidity that arises due to their ability to accurately forecast its flexibility potential only in the short-term; b) improved grid forecasts closer to real-time mitigate volume risk for network operators. Nevertheless, long-term contracts are expected to persist as a means of securing reliability and integrating flexibility into the long-term expansion of networks.

The focus of this work is on the short-term horizon, drawing inspiration from successful System Operator pilots such as the CoordiNet project Swedish pilot [25], sthlmflex project with NODES platform [23], and the Redispatch 2.0 in Germany [26]. Additionally, insights are drawn from market operator initiatives such as Enera Flexmarkt, operated by EPEX SPOT [23], and OMIE with the IREMEL platform [4]. It is important to note that this methodology extends beyond flexibility market procurement. It can also be applied for defining curtailment signals, particularly for RES with non-firm connection contracts (like Enedis in France [14]).

For this work, the following assumptions were made. Firstly, day-ahead and intraday flexibility markets occur before each energy trading session, which aligns with the flexibility market platforms offered by NODES, GOPACS, and Enera [4]. The market operator collects the bids submitted by FSP and provides this information to Transmission System Operator (TSO) and/or DSO. The flexibility needs can be revised in the intraday sessions with updated load and RES forecasts and changes in the flexibility band. Secondly, availability and dispatch payments are used for contracted flexibility. Thirdly, a pay-as-bid pricing method is considered for flexibility payment, widely recognized as the most common payment scheme [4].

The availability price is a mitigation measure to ensure enough flexibility volume in the market [4]. However, it is important to acknowledge that liquidity issues related to flexibility markets can still arise. For instance, the price cap established through the long-term analysis [24] may significantly limit revenue opportunities for Flexibility Service Providers (FSPs) in technical constraints management markets, particularly for those already involved in more lucrative markets such as wholesale energy trading and frequency control. While addressing liquidity problems falls beyond the scope of this work, the proposed methodology offers valuable contributions to mitigate this issue. Firstly, the implementation of *second-level forecasting* effectively reduces the overall cost associated with flexibility use, as shown in section V. This approach minimizes expenses and discourages an open-book mentality in

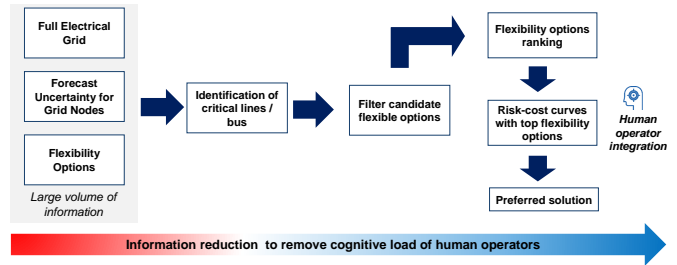


Fig. 1: Building blocks of the risk-aware flexibility procurement

the process of flexibility procurement. Secondly, the proposed methodology also serves as a valuable tool to communicate the potential risk of flexibility scarcity to the operator, as illustrated in section V-E.

Furthermore, without loss of generality, this work considers hourly market intervals and three forecasting moments: i) day-ahead forecast (for $D+1$) with numerical weather predictions (NWP) generated at 0h00, i.e., between $t + 24|t$ and $t + 48|t$; ii) updated NWP data at 12h00, i.e., between $t + 12|t$ and $t + 36|t$; iii) 2-hour forecast before delivery, which can be used for intraday participation, i.e., $t + 2|t$.

Finally, the study and integration of various TSO-DSO coordination schemes are outside the scope of this work. However, it is important to note that information from both TSOs and DSOs can be effectively integrated into the proposed methodology, as elaborated in the subsequent section.

B. Methodology

Fig. 1 presents the building blocks of the proposed methodology. It departs from a large volume of information, namely the full electrical grid, RES/load uncertainty forecasts, and a set of flexibility options, i.e., dispatchable distributed generation, RES curtailment, demand response (DR), storage, network reconfiguration, on-load tap changer (OLTC) and capacitor banks/reactor shunts (CB/RS). Step by step, all this information is filtered and finally condensed into a risk-cost curve from where the human operator can select a preferred solution. This approach reduces the cognitive load of human operators and offers simplicity when analyzing the multiple flexibility options under uncertainty. Moreover, it enables simultaneous analysis of multiple lines/buses with a high probability of technical problems. The load and RES forecast uncertainty is represented by random vectors (or scenarios) that can be generated by physical or statistical approaches and capture the spatial dependency of forecast errors.

The first step focuses on running a power flow for each scenario, and computing, for each node and branch, the probability of having an over/under-voltage and congestion issue. The operator can set a minimum probability to select and analyze a reduced set of critical buses/lines. For this subset of elements, sensitivity indices (see section III-A) relating active/reactive power and the line's current and node's voltage are computed. Then, in a second step, the flexibility options are filtered based on a sensitivity index minimum threshold to identify the most "relevant" options for a specific technical problem.

The third step includes performing a flexibility ranking with the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method [27] (section IV-A), considering a set of risk metrics that measure how effective a flexibility option is in solving a technical problem under uncertainty. It can also include probabilistic information (risk metrics) from the TSO or DSO associated with conflicting actions. Finally, the final step focuses on the design of a risk-cost curve (section IV-B) by combining a subset of flexibility options (e.g., 3 to 5 options) selected by the operator and considering only the non-dominated solutions. The flexibility capacity of these options may also be constrained in the curves due to conflicting actions arising from the activation of flexibility by both DSO and TSO. The risk can be represented by different metrics, and without loss of generality, this work uses the probability of having a congestion or under/over-voltage problem. To obtain a preferred solution, we consider two decision-making paradigms: one based on the maximum risk threshold and the other based on a trade-off value (section IV), where the risk threshold and trade-off values are conditioned by the stakes level (i.e., risk of cascading failure) and operator's preferences. Note that this relation between risk and stakes borrows the fundamental idea behind confidence-based decision-making theory [28].

The methodology above does not provide information about the best moment to book flexibility, considering that i) forecast uncertainty will decrease with time and ii) the flexibility availability price might increase with a shorter notification time. Thus, a *time-to-decide* problem is formulated based on the *second-level forecasting* concept (section III-D), where the goal is to forecast how the RES and load uncertainty forecasts change over time, as new information is available. The term *second-level forecast* comes from the fact that we are, in essence, computing a forecast of forecast uncertainty. For the flexibility price, it is assumed that the SO has a forecast or that the price offers with different notification times are available in advance.

The *second-level forecast* of conditional quantile $\hat{q}_{t+k|t}^\alpha$, Eq. 1, generated with information available at time instant t (forecast launch time) for time interval $t+k$, with nominal proportion α ($0 < \alpha < 1$), and a forecast update rate of z time intervals, corresponds to the conditional expected value of quantile $\hat{q}_{n,t+k|t+z}^\alpha$ of the next forecast (i.e., generated at $t+z$) for the same time interval $t+k$. f is the *second-level forecasting* model for each conditional quantile α , and uses as covariates i) quantile forecasts $\hat{q}_{t+k|t}^\alpha$ generated at t and ii) engineered features that explain the level of uncertainty in the power forecast generated at t and exogenous variables such as NWP, both denoted by $X_{t+k|t}^\alpha$ with l variables, plus a random shock $e_{t+k|t}$.

$$\begin{aligned}\hat{q}_{t+k|t}^\alpha &= \mathbb{E}[\hat{q}_{t+k|t+z}^\alpha | \hat{q}_{t+k|t}^\alpha, \mathbf{X}_{t+k|t}^\alpha] \\ &= f\left(\hat{q}_{t+k|t}^\alpha, \hat{x}_{t+k|t}^1, \dots, \hat{x}_{t+k|t}^l\right) + e_{t+k|t}\end{aligned}\quad (1)$$

Fig. 2 depicts the integration of *second-level forecasting* into the *time-to-decide* problem, considering: i) forecast launched at 0h00 (t) for lead-time $t+30$ (day-ahead), and ii) two *second-level forecasts* for the forecast that will be

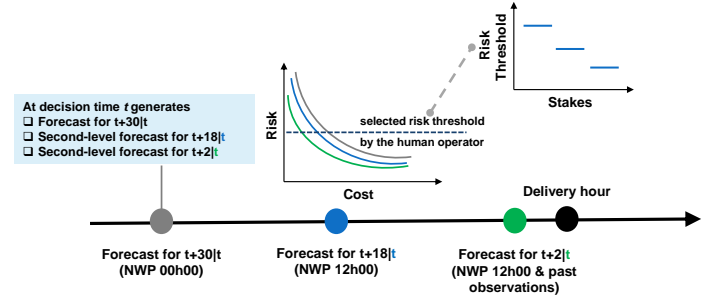


Fig. 2: *Time-to-decide* framework and risk-stakes relation. The integration of *second-level forecasts* into the *time-to-decide* problem generates risk-cost curves for a) forecast launched at 0h00 for lead-time $t+30|t$, and b) two *second-level forecasts*, one for the $t+18|t$ forecast (NWP updated at 12h00), and another for the $t+2|t$ forecast. The preferred solution for the human operator is selected from a risk threshold conditioned by the risk-stakes relation

generated 12 hours later with NWP updated at 12h00 (i.e., $\hat{q}_{t+30|t}^\alpha = \mathbb{E}[\hat{q}_{t+30|t+12}^\alpha | \hat{q}_{t+30|t}^\alpha, \mathbf{X}_{t+30|t}^\alpha]$), and another for the forecast that will be generated 2-hours before delivery time (i.e., $\hat{q}_{t+30|t}^\alpha = \mathbb{E}[\hat{q}_{t+30|t+28}^\alpha | \hat{q}_{t+30|t}^\alpha, \mathbf{X}_{t+30|t}^\alpha]$). This leads to three risk-cost curves that inform the operator about the possible outcome of waiting for the next forecast to book flexibility, each one corresponding to three different forecasting launch times for the “delivery hour”, $t+30|t$, $t+18|t$, and $t+2|t$. In this illustrative example, the preferred option would be to wait for the next forecast since the curve obtained with the *second-level forecast* for $t+2|t$ indicates a lower flexibility cost for the same risk level.

Finally, it is essential to mention that contingencies were not considered. However, the methodology can be easily extended to include information about the probability of contingencies in the risk modeling and combined with the scenarios.

III. KNOWLEDGE CONSTRUCTION FOR FLEXIBILITY PROCUREMENT

A. Sensitivity indices (SI)

1) *Analytical method*: The Y_{bus} compound matrix method [29] was used for the analytical derivation of node voltages and line currents SI. Using the complete admittance matrix of the network, this method takes advantage of the sparsity in the admittance matrix. It achieves computational efficiency surpassing that of Jacobian-based methods. Moreover, it applies to networks with any number of slack buses. For topology reconfiguration, the impedance (Z_{bus}) matrix method was used to obtain SI (distribution factors) due to its linearity and computational efficiency [30].

2) *Machine learning proxy*: The SI vary with the grid operating conditions, meaning each scenario requires an analytical calculation. To decrease the computational time due to its fast inference time, the gradient boosting trees (GBT) model is used as a proxy to extract functional knowledge relating SI and grid operating conditions. Firstly, a large set of grid operating scenarios is generated with the NORTA (NORmal To Anything) method [31]. Then, the SI are analytically computed for this generated set of grid operating scenarios, and the

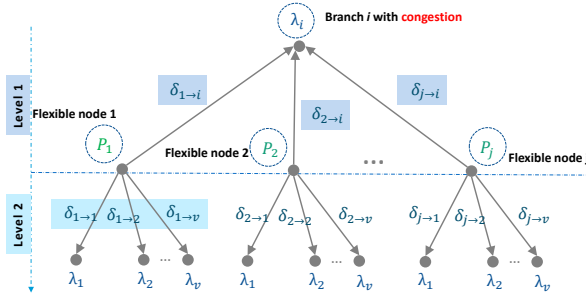


Fig. 3: Causality tree for a congested line i and flexibility options in nodes j (Level 1). Level 2 corresponds to the inverse impact in other grid elements by applying the Level 1 actions

GBT is fitted over this data using active and reactive nodal power injection as input features. To keep low the number of input variables while leveraging on the localized impact of active/reactive power nodal injections, the Spearman rank correlation coefficient was applied to measure the correlation between the time series of nodal power injection and the SI. Only the nodal injections exhibiting a correlation above 0.5 were selected as input features.

The GBT model is then used to derive SI $\delta_{i \rightarrow j, s}^{t+k}$ (Eq. 2) for the forecasted scenario s , and for grid element i (i.e., line/bus) and flexible resource j , and using as input a matrix of forecasted active $\hat{\mathbf{P}}_s^{t+k}$ and reactive $\hat{\mathbf{Q}}_s^{t+k}$ power nodal injections for lead-time $t+k$. The mean absolute percentage error of estimated SI with GBT was less than 1%.

$$\hat{\delta}_{j \rightarrow i, s}^{t+k} = f(\hat{\mathbf{P}}_s^{t+k}, \hat{\mathbf{Q}}_s^{t+k}) \quad (2)$$

The analytical method is only used for generating the training set of the GBT proxy, using network operating scenarios generated by the NORTA method. Subsequently, the GBT model is used in the remaining steps of the methodology.

B. Interpretation from sensitivity indices

The SI relate the rate-of-change in nodal voltage/line current as a linear function of active/reactive power modulation, variation in CB/RS elements, or OLTC tap position. This information could assist the human operator in figuring out the causal relationship between the technical issue and the flexibility options. Fig. 3 depicts a causality tree divided into two levels (unrelated to the network topology), which is helpful to guide the human operator analysis to the most suitable flexibility options. In Level 1, the SI represent the influence of each flexibility option on reducing the overload in line i via downward/upward (active power) action over the current operating point of each bus j . Since the activation of flexibility can have an inverse impact on other grid elements (e.g., cause congestion in another line), the SI in Level 2 represent the negative (inverse) impact on other grid elements, which can be used to constrain the activated flexibility in each node j .

From the SI and scenarios, it is possible to compute the contribution C_j of each grid node (load or generator) j forecast uncertainty to the forecasted technical problem in grid element

i , via the average deviation of a set of S scenarios around the point forecast $\hat{y}_{t+k|t,j}^*$ weighted by the SI:

$$C_j = \frac{1}{S} \sum_{s=1}^S \frac{\hat{y}_{t+k|t,j}^* - \hat{y}_{t+k|t,j}^s}{\hat{y}_{t+k|t,j}^*} \cdot \delta_{j \rightarrow i}^s \quad (3)$$

where $\delta_{j \rightarrow i}$ is the sensitivity of resource j to grid element i , and for scenario s .

C. Flexibility modeling

1) *Active and reactive power nodal flexibility*: Comprises a FSP aggregating DER or individual RES power plants, DR, and storage system per grid node, providing flexibility via upward/downward actions around their operating point. The FSP can use advanced aggregation functions, such as [32], to effectively manage and optimize the heterogenous DER, considering technical constraints, market dynamics, operational objectives, and different asset owners.

The flexibility value (FV), Eq. 4, represents the amount of active/reactive power that is computed per flexible resource j , based on $\delta_{j \rightarrow i}^s$, to solve a technical problem considering the max/min operational limit λ^{lim} (e.g., voltage or current limits) and the forecasted operating conditions λ_i^s for each scenario s , where i means the element with a technical issue. The suggested upward ($Flex_j^{up}$) and downward flexibility margin ($Flex_j^{down}$) in the flexibility market is an upper limit for $FV_{j \rightarrow i}^s$.

$$FV_{j \rightarrow i}^s = \frac{\lambda^{lim} - \lambda_i^s}{\delta_{j \rightarrow i}^s} \quad (4)$$

As illustrated in Fig. 3, a specific flexibility action may create technical issues in other grid elements with opposite SI signs. Thus, Eq. 5 is applied to limit the $FV_{j \rightarrow i}^s$ value, conditioned by the operational limits of the other grid elements v .

$$\begin{aligned} FV_{j \rightarrow i}^s = & \\ \text{sgn}(FV_{j \rightarrow i}^s) \cdot \min\{|FV_{j \rightarrow i}^s|, \min_{v \in V, v \neq i} \{|\frac{\lambda^{max} - \lambda_v^s}{\delta_{j \rightarrow v}^s}|\}\}, & \\ \text{if: } \text{sgn}(\delta_{j \rightarrow i}^s) \neq \text{sgn}(\delta_{j \rightarrow v}^s) & \end{aligned} \quad (5)$$

In situations where conflicting actions between TSO and DSO may arise during the activation of flexibility, Eq. 5 can be applied to limit the $FV_{j \rightarrow i}^s$ value. For example, suppose activating a flexible resource within the distribution network leads to technical issues in the transmission network. In that case, the TSO can communicate the parameters of Eq. 5 to enable the DSO to compute the corresponding limitation. Alternatively, the TSO can simply communicate the limitation percentage applicable to that specific resource.

Reaching the lower limit λ^{min} for elements with the same SI sign is also possible where the elements are forced to a minimum operational limit (i.e., bus voltage should be above 0.9 p.u.). In this case, Eq. 6 is applied.

$$\begin{aligned} FV_{j \rightarrow i}^s = & \\ \text{sgn}(FV_{j \rightarrow i}^s) \cdot \min\{|FV_{j \rightarrow i}^s|, \min_{v \in V, v \neq i} \{|\frac{\lambda^{min} - \lambda_v^s}{\delta_{j \rightarrow v}^s}|\}\}, & \\ \text{if: } \text{sgn}(\delta_{j \rightarrow i}^s) = \text{sgn}(\delta_{j \rightarrow v}^s) & \end{aligned} \quad (6)$$

Moreover, the activation of $FV_{j \rightarrow i}^s$ also creates a change in the slack bus active and/or reactive power injection (in the opposite direction of the flexibility use), which can lead to technical problems in grid elements with the same SI sign of $\delta_{j \rightarrow i}^s$. Thus, Eq. 7 is applied to limit the $FV_{j \rightarrow i}^s$ avoiding a second order technical problem due to the slack bus (sb) operating point variation.

$$FV_{j \rightarrow i}^s = \text{sgn}(FV_{j \rightarrow i}^s) \cdot \min\{|FV_{j \rightarrow i}^s|, \min_{v \in V, sb \in SB} \left\{ -\frac{\lambda_{max} - \lambda_v^s}{\delta_{sb \rightarrow v}^s} \right\}\}, \text{Eq. 10:}$$

$$\text{if: } \text{sgn}(\delta_{j \rightarrow i}^s) = \text{sgn}(\delta_{sb \rightarrow v}^s) \quad (7)$$

where V is the line and bus elements set, and SB is the set of slack buses in the system. The same constraint in Eq. 6 is applied for the slack bus but conditioned to grid elements with the opposite SI sign of $\delta_{j \rightarrow i}^s$ due to a reverse reaction of the slack bus to the flexibility use.

2) *Active power redispatch*: Combined upward/downward action between two (or more) flexible resources/nodes, such that total active power feed-in remains virtually unchanged, but the congestion is removed. Note that, currently, redispatch actions are primarily employed by TSOs and are included in our proposed methodology for completeness. The application of such actions in distribution networks is still a subject of ongoing discussion, as highlighted in [33]. However, initiatives such as redispatch 2.0 and 3.0 in Germany [26] have the potential to extend redispatch actions to distribution networks.

In this case, the individual SI in Eq. 4 are replaced by the ones computed by Eq. 8 for several n flexibility resources (i.e., j_1, j_2, \dots, j_n) participating in redispatch, where $b_{j \rightarrow i}^s$ is an indicator function equal to +1 if the action is upward and -1 if downward.

$$\delta_{\{j_1, j_2, \dots, j_n\} \rightarrow i}^s = \sum_{j=j_1}^{j=j_n} b_{j \rightarrow i}^s \cdot \delta_{j \rightarrow i}^s \quad (8)$$

To avoid technical issues in other elements (lines or buses) Eq. 5-7) are checked for all units participating in re-dispatch.

3) *Reactive power flexibility from CB/RS*: The FV for shunt elements follows Eq. 4 and moves on discrete steps determined by the steps of the CB/RS unit as formulated in Eq. 9. The FV is communicated to the human operator regarding step changes in shunt elements.

$$FV_{j \rightarrow i}^s \in \{\pm Q_j^u, \pm 2Q_j^u, \dots, \pm stp_j^{max} \cdot Q_j^u\} \quad (9)$$

where stp_j^{max} is an integer parameter representing the maximum allowed step-change in CB/RS and Q_j^u is the reactive power per step change in shunt element j .

4) *OLTC*: : in this case, the FV in Eq. 4 will be equal to $tap^{[\%]} \cdot tpos$, where $tpos$ and $tap^{[\%]}$ mean tap-position and tap-percentage (i.e., percentage of voltage adjustment versus one-step change in tap-position). Thus, Eq. 4 is reformulated to Eq. 10, where λ_i^s is voltage in bus i and scenario s .

$$FV_{j \rightarrow i}^s = tap^{[\%]} \cdot tpos^s = \frac{\lambda^{lim} - \lambda_i^s}{\delta_{j \rightarrow i}^s} \quad (10)$$

Moreover, the goal is to determine the required voltage set-point that will be automatically translated to a tap position by the OLTC, using the relation from Eq. 11.

$$V(\text{set} - \text{point})_j^s = (1 + tap^{[\%]} \cdot tpos^s) \cdot V(\text{set} - \text{point})_j^{ne} \quad (11)$$

In Eq. 11, if the tap position ($tpos^s$) remains zero, the voltage will be unchanged and equal to the voltage set-point of OLTC j in neutral tap position $V(\text{set} - \text{point})_j^{ne}$ (i.e., zero tap position). Then, Eq. 11 could be revised as Eq. 12 using

$$V(\text{set} - \text{point})_j^s = (1 + FV_{j \rightarrow i}^s) \cdot V(\text{set} - \text{point})_j^{ne} \quad (12)$$

Like other flexibilities, the FV should be limited by the operating margins of other grid buses using an equation analogous to Eq. 5. Similarly, the limitation of flexibility capacity can also arise due to conflicting actions between TSO and DSO, as explained earlier.

5) *Network topology*: A combination of switching actions proposed to change the network topology while ensuring it remains a traceable-connected graph. The candidate switching actions are selected based on their *area-of-effectiveness*, defined as the minimum path between two-side buses of the switched line using graph theory. This approach avoids simulating all possible switching actions. Therefore, the closing/opening actions are determined as follows: a) the closing action is applied to normally open lines whose *area-of-effectiveness* includes the congested line; b) the opening action is applied to normally closed lines in the *area-of-effectiveness* of the selected closing action. It was assumed that the opening action only applies when preceded by a closing action within the same area to prevent additional technical issues. As a result, the proposed actions can include a single closing action, sequential closing actions, and sequential closing and opening actions. This type of flexibility is considered in the decision-aid phase of section IV together with the other flexibilities.

D. Second-level forecasting

The *second-level forecasting* concept builds upon a set of computed probabilistic forecasts (set of quantiles) generated with any forecasting algorithm. In this work, we used a GBT model with feature engineering [34], with truncated generalized Pareto distribution for the distribution's tails [35]. This GBT model was extended with additional input variables for *second-level forecasting*, namely variables engineered with the quantiles from the first-level forecast, i.e., forecast for the day $D + 1$ with NWP data generated at 0h00 of day D .

A second model was an Encoder-Decoder Artificial Neural Network (ED-ANN) [36]. The Encoder comprises two layers, which are Long Short-Term Memory (LSTM) network layers. The LSTM has feedback connections and can model temporal or sequential information. The Decoder comprises three internal layers: one LSTM and two fully connected layers, with linear activation function. The Decoder produces the output iteratively, and the hidden state produced at each forecast step generates the next forecast. In the training phase, teacher forcing was used, i.e., using the actual first-level forecast values as input for training the Decoder.

IV. DECISION-AID PHASE

A. Flexibility ranking

After calculating the $FV_{j \rightarrow i}^s$ (see section III), the next step is ranking the individual flexibility options according to their *effectiveness* in solving the technical problem under forecast uncertainty. Firstly, the line loading/bus voltage after flexibility use is computed for each scenario:

$$\lambda_i^{\prime s} = FV_{j \rightarrow i}^s \cdot \delta_{j \rightarrow i}^s + \lambda_i^s \quad (13)$$

where λ_i^s and $\lambda_i^{\prime s}$ are the line loading/bus voltage of element i before and after flexibility activation.

Secondly, for each scenario, s , the severity function from [37] associated with line congestion/voltage violation (see Eq. 14) is applied to the value of $\lambda_i^{\prime s}$, where $\lambda_i^{\%}$ is the loading percentage of the line i (an analogous equation is applicable for voltage problems).

$$Severity_i = 0.1 \cdot (\lambda_i^{\%} - 90) \quad (14)$$

The flexibility cost corresponding to $FV_{j \rightarrow i}^s$ is also computed considering the flexibility bid price. The cost of the system operator assets was computed with the methodology from [38], described in Appendix A.

It is important to underline that the goal here is to rank each flexibility option. Thus, cost and severity are computed for each scenario and flexibility option. Then, risk metrics are computed from the severity and flexibility cost, namely: expected value and value-at-risk (VaR) of the flexibility cost, expected value and VaR of the severity, probability of over/under-voltage, or congestion. In situations involving conflicting actions between TSO and DSO, an alternative approach to constrain the $FV_{j \rightarrow i}^s$ in Eq. 5 (as explained in section III-C) is to incorporate risk metrics that quantify the potential impact on the TSO or DSO networks. This can include metrics such as the probability of congestion occurring in an upstream line within the TSO network.

Note that these metrics are computed for all flexibility options described in section III-C, including network reconfiguration actions (whose impact is estimated with the Z_{bus} matrix – see section III-A1). Then, the TOPSIS method is applied to rank the flexibility options (i.e., set of alternatives), using the risk metrics as criteria. There is no differentiation between flexibility types during the ranking process. For instance, if a network switching action ranks among the top solutions, it will be included in the set of actions used to construct the risk-cost curves discussed in the following subsection.

B. Risk-cost curves

The top three to five flexibility options are selected for combination following the flexibility ranking. All the possible combinations, including binary actions and actions with discrete and continuous values, are considered. Only the non-dominated ones create the risk-cost curve among all these possible combinations.

This approach is applied to the uncertainty forecast available in the present time and to the *second-level forecasts* generated with the methodology described in section III-D, which will result in multiple risk-cost curves (one for each forecast).

C. Decision-making paradigms

The final step (*decision*) requires the involvement of a human decision-maker to select a preferred solution from the risk-cost curve(s). Two alternative decision-making paradigms are considered:

1) *Maximum risk-threshold*: Sets a maximum threshold for the risk level conditioned by the decision stakes. In this work, the stakes are directly linked to the risk of cascading failures in the grid if the technical problem is a congested line. A similar concept can be derived for voltage problems, e.g., propagation effects of over-voltage in the protection system. The cascading risk is computed by simulating (using the Z_{bus} method, in each scenario s) the impact on all the grid lines of tripping the congested line i and computing the expected number of lines tripping. This expected value is normalized between 0 and a maximum pre-defined value for the cascading simulation and re-scaled into a scale of integer numbers that is easy to analyze by a human operator (e.g., between 1 and 10).

It is necessary to find, via interaction with the human decision-maker, a functional relation between stakes (risk of cascading) and maximum risk threshold so that stakes condition the risk level in each decision. A simple method to find this functional relation is a direct rating (e.g., the decision-maker is asked to state the maximum risk for stakes equal to 10). Still, other methods like direct mid-point can be used [39].

2) *Risk-cost trade-off*: The preferred solution is the one with the minimum equivalent cost, $Eqcost$ (Eq. 15), obtained by considering a trade-off value μ between risk and cost for each solution p in the curve. The trade-off value can be determined by interaction with the decision-maker, such as indifference judgments [39], and estimated for each stake level, creating a functional relation between trade-off value and stakes. In this case, the higher the stakes, the higher the trade-off value, i.e., the decision-maker is willing to pay more to decrease risk.

$$Eqcost_p = cost_p + \mu \cdot risk_p \quad (15)$$

V. NUMERICAL RESULTS

A. Case-study description

The 20 kV Oberrhein MV network supplied by two 25 MVA HV/MV substations¹ was used with some modifications to increase RES penetration and create technical problems. The network supplies 141 MV/LV (secondary) substations and 61.86 MW loads (peak power) through four MV feeders. The topology is meshed but is operated as a radial grid. It contains 147 consumers, nine WPP, four CHP units, and three storage systems. The secondary substation load measurements were taken from the Iowa Distribution Test System [40], and the wind power measurements are from WPP in France (the location cannot be disclosed for confidentiality reasons). The NWP are from the ECMWF High-Resolution Forecast model. Without loss of generality, the uncertainty forecast has been applied only for WPP, and a perfect forecast was used for the

¹https://pandapower.readthedocs.io/en/v2.0.0/networks/mv_oberrhein.html

TABLE I: Functional relation between stakes and risk threshold/trade-off value for different DM

Decision-making approach	Stakes (ρ) range		
	$0 \leq \rho \leq a^*$	$a < \rho \leq 7$	$7 < \rho \leq 10$
DM A: Maximum risk threshold	10	6	3
DM B: Maximum risk threshold	20	15	10
DM C: Maximum risk threshold	25	20	20
DM D: Risk-cost trade-off	30	50	70
DM E: Risk-cost trade-off	70	90	110

* Parameter a for DM A is equal to 3, for DM B, DM D and DM E is equal to 4, and for DM C is equal to 2.

substation load and CHP. The training period is from 2019-04-01 to 2020-03-30, and the testing period is six months, starting on 2020-04-01.

In terms of flexibility, the bid by generation units (i.e., WPP and CHP) was considered as 30% of the point forecast for upward/downward directions, while the upward direction for WPP was not considered. As assumed in [41], a shorter notification time can increase the flexibility price. Thus, the following was considered: 40% price increment for booking flexibility 12 hours later (assuming the present moment as the reference time) and 90% for booking two hours before the delivery time. A sensitivity analysis of these prices is presented in Section V-F. Without loss of generalization, it was assumed that availability and activation prices for flexibility are equal.

For DR, we considered a specific time schedule for its availability, and the maximum flexibility is 30% of the electrical energy consumption in a specific hour for upward/downward directions. In the case of storage units, flexibility was determined by the FSP as a function of battery state-of-charge and depth-of-discharge.

To evaluate the impact of different human decision-makers (DMs), different curves relating stakes and risk threshold/trade-off are considered in Table I.

The proposed methodology can address both congestion and voltage issues. However, the main focus was on congestion problems.

B. Benchmark Strategies and Evaluation Metrics

For benchmark, the following strategies were considered:

Time-to-decide (T2D): Novel concept proposed in this work where the operator decides i) book flexibility now or ii) wait for the next forecast update and book flexibility later.

Decision-now (DN): The operator, using the information from uncertainty forecasts, decides to book flexibility at the lowest (availability) cost based on forecasts generated with NWP generated at 0h00.

Deterministic 1 (D1): The operator decides, based on a deterministic forecast, to i) book flexibility now or ii) wait for the next forecast update and book flexibility later.

Deterministic 2 (D2): The operator always decides based on the very short-term forecast, i.e. $t + 2|t$.

The *probabilistic method* from [15].

The overall evaluation of the results considered two evaluation matrices: the classical confusion matrix and the cost-

TABLE II: Cost-loss matrix to evaluate the performance of the predictive grid management process

	Event occurred	Event did not occur
Action taken	Cell A: Rate-of-occurrence (h%) Flex cost (C) + Loss (L)	Cell B: Rate-of-occurrence (m%) Flex cost (C)
	Cell C: Rate-of-occurrence (f%) Loss (L)	Cell D: Rate-of-occurrence (c%) Zero cost

loss matrix (see Table II). The confusion matrix is used to evaluate the performance of each strategy in detecting congestion problems, from which F-score accuracy metrics are computed to measure the balance between the precision (i.e., dividing the true positives (TP) by anything that was classified as a positive) and recall (i.e., dividing the TP by anything that should have been classified as positive).

The cost-loss matrix draws inspiration from the threshold-based cost-loss analysis method described in [42] for evaluating weather forecasts. For this problem, the matrix was adapted to the following: flex cost (C) corresponds to the preventive actions cost (i.e., flexibility cost); loss (L) corresponds to real-time emergency action (i.e., load or generation curtail) to solve the congestion problem with a monetary cost corresponding to the value-of-lost-load (VoLL), considered to be 12000 €/MWh [43]. The table can be summarized into the summation of the matrix elements, each weighted by the percentage of occurrence, which is performance indicator γ :

$$\gamma = (C + L) \cdot h + C \cdot m + L \cdot f + 0 \cdot c \quad (16)$$

C. Second-level forecasting performance

1) *Input variables analysis*: The set of input variables is divided into five categories. Group I consists of variables computed from the NWP data (i.e., generated at 0h00 in day D), namely: the wind module and direction at 10 and 100 meters, averaged for the locations of the wind turbines in each wind power plant, lags and leads ($t \pm 3$) of these variables, as described in [34]. The first three components of the PCA decomposition of the NWP variables across all wind power plants (WPP) are also part of this group. Group II pertains to calendar variables, namely sine and cosine transformations of the hour of the day.

Group III are variables directly extracted from the first-level forecasts, namely the quantiles 5%, 95%, and the target quantile. Group IV is a proxy for the uncertainty level extracted from the first-level forecasts, namely: the difference between quantiles one position away from the target quantile (i.e., previous and following quantiles); the difference between quantiles two positions away from the target quantile; the average and standard deviation of the set of target and neighboring quantiles (two before and two after the target quantile); inter-quantile range.

Group V includes variables computed from a poor man's ensemble [44], given that we have a time horizon of up to 90 hours and 12 hours NWP updates. Using these variables did

TABLE III: List of input data variations for the *second-level* forecasting model

Name	Variable Groups	Modifications
Var1	I-IV	Include another set of variables, identical to Group I, but with NWP variables generated at 0h00 of day D-1. This set does not include lags and leads
Var2	I-IV	Same as Var1 but not using lags and leads in both Group I and the identical set added
Var3	I-IV	Same as Var1, using lags and leads in both Group I and the identical set added
Var4	I-III	
Var5	I-V	
Var6	I-V	Group V is modified using averages and std. dev. of the differences in NWP consecutive runs

not yield any improvement in almost all quantiles. In fact, on average, the model performed was worse than the best model, i.e., -0.17% improvement in the *second-level forecast* for the 12h forecast, and -0.03% for the 2-hours ahead *second-level forecast*. The meteo-risk index proposed in [44] reflects the spread of the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble forecasts was also tested as an input variable. The improvement was below 2.9%, which does not justify the cost associated with this data. Different variants of input variables were tested, as shown in Table III.

The *second-level forecasting* skill is evaluated with the Mean Squared Error (MSE) considering the difference between forecasted quantile ($\hat{q}_{t+k|t+z}^\alpha$) and *second-level forecast* of the same quantile ($\hat{q}_{t+k|t}^\alpha$). Moreover, a naive benchmark model where $\hat{q}_{t+k|t}^\alpha = \hat{q}_{t+k|t}^\alpha$ was used, assuming that there is no variation in the quantile forecasts with an information update. The MSE improvement over the benchmark model was computed for each model. As mentioned in Section II, two *second-level forecasts* are considered: *second-level forecast* for forecast updated with NWP generated at 12h00, and *second-level forecast* for two hours-ahead forecast. We used data from nine wind power plants (WPP) for our experiments. Data from April 2019 to March 2020 was used for model training, and from this portion, about 25% random samples were selected for validation. The testing period spanned from April to September 2020. Note that the input and target quantiles of the training period have been produced through k-fold cross-validation.

Table IV presents the mean (overall WPP, quantiles 5% to 95% with 5% increments) improvement over the naive model for different model variations from Table III. The best performance is from model variation Var1. This combination of variables is used to compare GBT and ED-ANN.

2) *Model comparison*: The results, averaged by WPP and including extreme quantiles (lower than 5% and higher than 95%), are shown in Table V. Starting with the *second-level forecast* for the 12h00 forecast launch time in day D , we observe that ED-ANN shows results comparable with the GBT, although the improvement is higher for the former in five WPP. For the *second-level forecast* for two hours-ahead forecast,

TABLE IV: Mean percentage improvement (per WPP) over the naive model for different model variations.

Second-level forecast	Model Variations					
	Var1	Var2	Var3	Var4	Var5	Var6
NWP generated at 12h00	9.17	7.27	9.36	9.11	9.10	8.39
2-hours ahead forecast	14.06	13.69	12.42	11.93	14.08	11.46

TABLE V: Improvement over naive model per WPP

Second-level forecast Model	WPP								
	0	1	2	3	4	5	6	7	8
12h forecast									
GBT	17.0	23.0	18.5	25.8	17.5	18.5	15.4	15.1	14.3
ED-ANN	13.7	25.2	18.6	26.1	18.9	16.5	15.8	13.0	11.9
2-hours ahead forecast									
GBT	19.4	32.7	19.9	16.3	18.7	22.6	17.9	13.6	23.2
ED-ANN	24.9	31.2	22.8	21.5	20.9	26.7	21.5	16.2	23.1

the ED-ANN outperforms the GBT model for seven WPP. Both improve on the naive model, showing the potential for exploring the *second-level forecasting* concept, in particular, the ED-ANN architecture.

D. Illustrative example

This subsection presents an example for line 155 on August 13, at 17h00, to demonstrate the potential of the *time-to-decide* approach in reducing false congestion alarms since the observed (real) loading of this line was 37.6%. The forecast launched at 0h00 for day D+1 (i.e., $t + 42|t$) estimated a probability of congestion equal to 39.6%. In contrast, the two *second-level forecasts* (for D+1 with NWP at 12h00, and for $t + 2$) show a decrease in the probability, 23.6%, and 16.6% correspondingly, which might indicate a false congestion alarm.

Fig. 4 illustrates the contribution of the nine WPP (C_j in Eq. 3), depicted through gradient color, for the false congestion alarm in line 155. In this case, the highest contribution (i.e., 36.7%) is from the WPP at bus 51 (where line 155 is connected), which has a higher forecast absolute error for that time interval (47.26% of rated power), followed by WPP connected to bus 31 with a contribution of 23.87% and an absolute forecast error of 52.53%.

Fig. 5 depicts the line loading's probability density function (PDF) for the three different forecasts obtained with the kernel density estimation over the scenarios samples. The point forecast (100.3% of the line loading) and observed value (37.6% of the loading) are also presented.

The risk vs cost curves for each forecast are depicted in Fig. 6. Using the risk threshold-stakes curve in Table I, and since, in this case, the stakes are equal to 6, the risk threshold is equal to 6% (in DM A) and 15% (in DM B). For DM A, book flexibility only two hours in advance ($t+2|t$) is preferred. Note that the stakes for the two *second-level forecasts* are equal to 4, which leads to a risk threshold of 6% for DM A. Therefore, the operator should wait for the upcoming forecasts and not book any flexibility. Then, 12 hours later, the operator

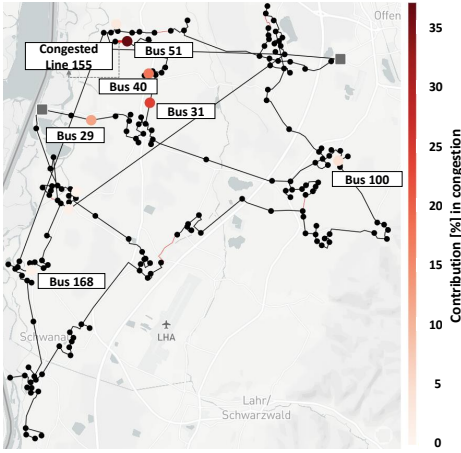


Fig. 4: Contribution (C) of RES power plants forecast uncertainty to the forecasted probability of congestion in one line

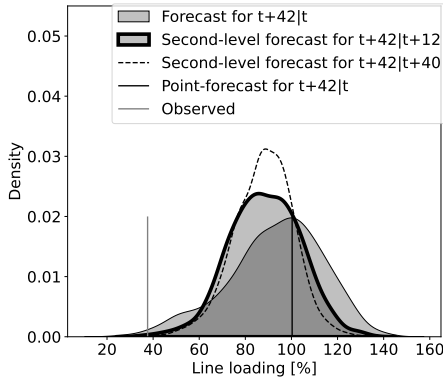


Fig. 5: PDF of the line loading obtained with day-ahead forecast and *second-level forecasts*

receives an updated forecast for the same hour (now lead-time $t + 30|t$) and a second *second-level forecast* for $t + 30|t + 28$. Due to the lower total flexibility cost in the *second-level forecast* for $t + 30|t + 28$, the operator will postpone the flexibility booking for the last forecast in $t + 2|t$. The very short-term forecast launched at $t + 2|t$ estimates a near zero probability of congestion, as depicted in Fig. 7, meaning that no flexibility is necessary for that specific hour. The same procedure could be followed for the risk level of DM B, and it would also reach this point where flexibility is not used. Therefore, the proposed *time-to-decide* methodology can result in a reduction of the flexibility cost and false alarms through the use of *second-level forecasts*.

E. Practical issues

1) *Flexibility scarcity*: Flexibility scarcity events might occur (e.g., due to low flexibility market liquidity) and can be detected in advance by the proposed method. For instance, on April 14, at 11h00, the observed loading at line 49 was 124.47%. This congestion was detected with the forecast for $t + 35|t$ with a probability of 89.9%. The risk-cost curve for the $t + 35|t$ forecast is depicted in Fig. 8. The set of actions

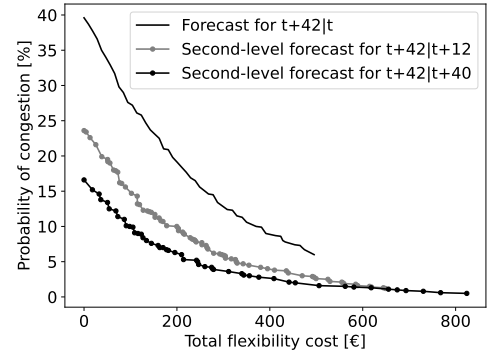


Fig. 6: Risk-cost curves for the forecast and two *second-level forecasts* launched at 0h00

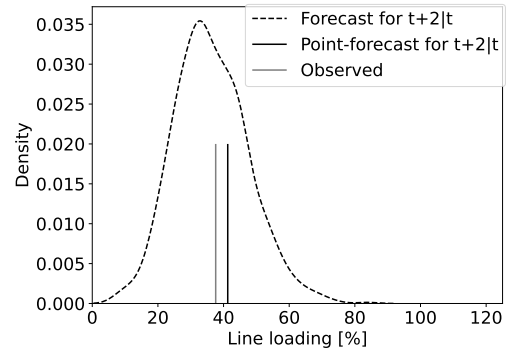


Fig. 7: PDF of line loading of forecast launched at $t + 2|t$

for the lowest reachable risk (i.e., 79.6%) are active power curtailment of 30% in WPP and CHP unit of bus 29 and WPP of bus 51 together with 20% in buses 43 and 117. However, these actions are insufficient for solving this technical issue or reaching an acceptable risk threshold for a decision-maker. In this case, the congestion problem will occur, leading to real-time control actions (e.g., RES curtailment) and a monetary loss of 329.76€.

2) *Conflicting action between TSO and DSO*: In this illustrative example, line 155 on July 5 at 04h00, a solution at the distribution level creates transmission-level problems.

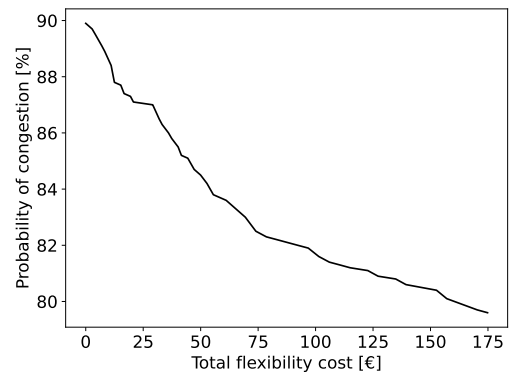


Fig. 8: Risk-cost curve for a case with flexibility scarcity

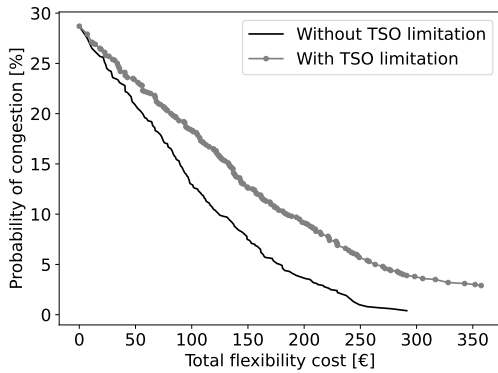


Fig. 9: Risk-cost curves for the cases with and without TSO limitation in the flexibility value (FV)

The TSO has two possibilities to integrate this information: a) incorporating risk metrics into the flexibility ranking that quantify the potential impact on the TSO network (as explained in section IV-A); b) imposing a limit on the $FV_{j \rightarrow i}^s$ value to prevent technical problems in the transmission network (as explained in section III-C).

To illustrate the first possibility, Table VI presents the modification in the flexibility ranking by incorporating the probability of congestion in the transmission network (referred to as “TSO risk” in the Table for simplicity) and the set of top flexible resources. Using this approach, the information from the TSO would be used to influence the flexibility ranking and select the actions with a lower probability of creating technical problems in the upstream network. This implies that flexibility options with higher costs may be placed in the top-ranked positions because of their capability to avoid technical issues in the transmission network. For the second possibility, Fig. 9 depicts the risk-cost curves as an example for the $t + 28|t$ forecast, with and without TSO limitation to the $FV_{j \rightarrow i}^s$. The curve with TSO limitations shows that to maintain the same risk threshold (e.g., 5%), a higher cost for flexibility is necessary. As a result, constraints on flexibility due to conflicting TSO-DSO actions could result in expensive solutions.

TABLE VI: Flexibility ranking without and with TSO limitation

Without TSO limitation		With TSO limitation
Flexibility actions	Flexibility actions	TSO risk [%]
APC WPP @ bus 31	APC WPP @ bus 51	8.33
APC WPP @ bus 65	APC WPP @ bus 65	16
APC WPP @ bus 51	APC WPP @ bus 31	77.42
APC WPP @ bus 29	APC CHP unit @ bus 30	71.58
APC CHP unit @ bus 30	APC WPP @ bus 29	98.44

¹ APC stands for active power curtailment.

3) *Changes in flexibility level:* The forecasted congestion of line 155 on July 5 at 05h00 was selected for sensitivity analysis of the flexibility level (i.e., maximum flexibility per flexibility resource). Three flexibility levels were considered:

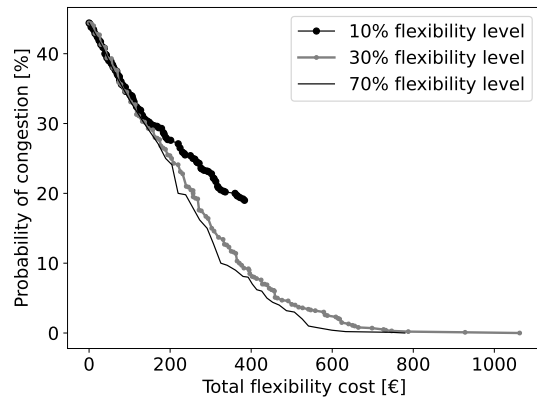


Fig. 10: Risk-cost curves for the $t+2|t$ forecast considering different flexibility levels

10%, 30% (which will be used in the remaining work), and 70% of the RES point forecast or load. The impact of this change on the risk-cost curve for $t+2|t$ is depicted in Fig. 10. It shows that reducing the flexibility level can prevent the curve from reaching a lower risk level. Moreover, for the same risk level, the total flexibility cost for a 70% flexibility level is lower since a larger volume of flexibility from top-ranked solutions is available. The result of the T2D strategy for different flexibility levels is presented in Table VII for DM A. In the 10% flexibility level, the risk-cost curve cannot reach the risk threshold defined by the DM, and real-time RES curtailment actions will be necessary to solve the congestion problem. Increasing the flexibility level to 30% and 70%, the congestion will be solved with the “reserved” flexibility. However, the cost of booking flexibility with a 70% flexibility level is lower. In both 30% and 70%, the solutions obtained by the T2D strategy outperform the DN strategy.

TABLE VII: Sensitivity analysis of the T2D strategy for different flexibility levels

Flexibility level	T2D approach			DN approach
	Flexibility cost [€]	Loss [€]	Total cost [€]	Total cost [€]
10%	78.9	945.5	1024.4	1024.4
30%	457.4	0	457.4	525.2
70%	426.1	0	426.1	520.3

F. Overall evaluation

An overall evaluation was conducted over six months from April to September 2020). All the simulations were run on a cloud-based virtual machine with AMD EPYC CPU with 8-core 2.94 GHz, 32 GB RAM, and Microsoft Windows 10 Pro. Table VIII presents the computational time of each step in the proposed method.

To evaluate the congestion detection performance of each strategy, the F1-score and F3-score are reported in Table IX. The T2D approach achieves the highest score, with DM A

TABLE VIII: Computational time of different steps in the proposed methodology

Step	Computational time (sec.)
Flexibility value calculation	64s (0.25s/scenario)
Flexibility ranking	0.016s
Risk-cost curve	47.59s
Decision-aid	0.14s

TABLE IX: Accuracy metrics for evaluating congestion detection performance and cost-loss matrix performance indicator- γ

Strategy	Precision	Recall	F1-Score	F3-Score	γ [M€]
D1	0.53	0.08	0.13	0.08	26.03
D2	0.53	0.03	0.06	0.03	27.03
Model from [15]	0.45	0.74	0.56	0.69	14.74
DN: DM A	0.45	0.71	0.55	0.67	5.05
T2D: DM A	0.56	0.79	0.66	0.76	3.98
DN: DM B	0.49	0.63	0.55	0.61	6.21
T2D: DM B	0.65	0.69	0.67	0.69	4.93
DN: DM C	0.50	0.60	0.54	0.59	6.63
T2D: DM C	0.66	0.66	0.66	0.66	5.27
DN: DM D	0.43	0.69	0.53	0.65	4.92
T2D: DM D	0.50	0.77	0.61	0.73	4.29
DN: DM E	0.44	0.69	0.53	0.65	4.76
T2D: DM E	0.48	0.78	0.59	0.73	3.72

(using a lower risk threshold) showing higher performance in F3-score, while DM B outperforms in F1-score. The deterministic strategies show the lowest score due to a lower rate of TP detection (and a high rate of FN cases). The probabilistic approaches overcome this limitation, and the rate of FP is reduced in the *T2D* strategy due to the information provided by the *second-level forecasts*. The benchmark model from [15] is outperformed by T2D in DM A, B, D, and E.

The results of the cost-loss matrix per DM strategies are presented in Fig. 11–12. The probabilistic strategies show a lower rate-of-occurrence and total cost in Cell C “action not taken, event occurred”. This shows their effectiveness in addressing overlooked congestion cases, in contrast to the deterministic strategies. The higher rate-of-occurrence and total cost in matrix row “action taken” (Cells A+B) are due to more congestion problems solved, mixed with some false congestion alarms (a “side-effect” of risk-based approaches).

Compared to the *DN* strategy, the *T2D* strategy shows a lower rate-of-occurrence and total cost for matrix Cell C “action not taken, event occurred” and Cell A “action taken, event occurred” since, with the use of *second-level forecasting*, the human operator may decide book flexibility now or postpone the decision to the next forecast. In this case, the strategy also benefits from the possibility of detecting overlooked congestion (i.e., reducing the FN cases). Moreover, the *T2D* strategy can significantly reduce the cost caused by false alarms, which is reflected in a higher rate-of-occurrence in Cell D “action not taken, event did not occur”. The possibility of selecting the flexibility booking moment by the *T2D* strategy can reduce the cost significantly, as explained in section V-D. The comparison between the three different DM, shows distinct results and that DM A with lower risk

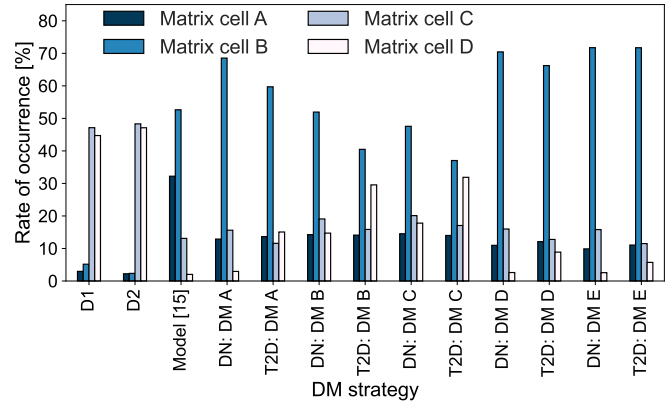


Fig. 11: Rate-of-occurrences for each cell of the cost-loss matrix.

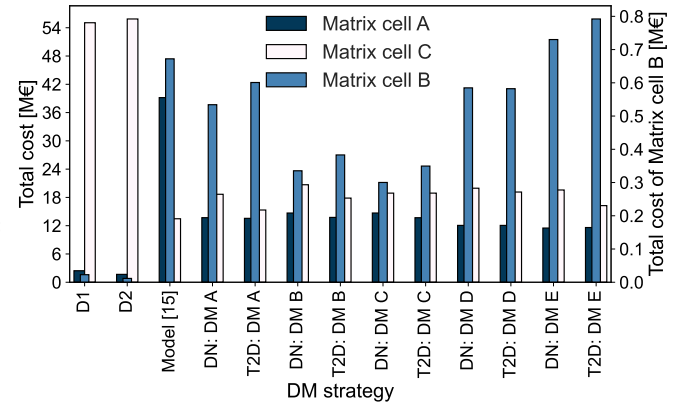


Fig. 12: Total cost for each cell of the cost-loss matrix.

thresholds (in comparison to DM B and C) can further reduce the loss of preventive management while paying more for the flexibility than the other two DM. The trade-off strategy (DM D and E) provides the DM with the ability to manage the balance between risk and cost effectively. For example, DM D was willing to pay 30€ and 70€ to achieve a 1% reduction in risk. This level of control over risk and cost is not possible with the risk-threshold strategy (DM A-C), where the DM can only determine the maximum acceptable level of risk without considering associated costs. Consequently, this limitation may result in the implementation of costly solutions to meet the risk threshold. In this case, DM D contracts more flexibility to solve the technical problems (higher rate-of-occurrence in Cell B “action taken, event did not occur”) compared to DM B and C, which have lower risk thresholds than DM A. DM E has higher trade-off value than DM D for each stakes level, which means that it is willing to pay more to decrease the risk. For instance, for DM E, this leads to a higher total cost in Cell B “action taken, event did not occur”, but also to a lower rate-of-occurrence than DM D.

The performance indicator γ was computed for each strategy, and results are presented in Table IX. These results show that the *D1* strategy that “reserves” flexibility in the moment with lower flexibility price, but higher uncertainty

(which leads to higher “reserved” flexibility), has 3.69% better performance than *D2* strategy. The *DN* strategy has 80.61%, 76.12%, 74.53%, 81.09% and 81.73 improvement over the *D1* strategy. The performance of the *T2D* strategy was evaluated concerning the *DN* strategy in each DM strategy, where an improvement of 21.08%, 20.72%, 20.45%, 12.78%, and 21.74% was obtained by DM A, DM B, DM C, DM D, and DM E, respectively. These results show the advantage of the proposed *T2D* strategy. Moreover, the benchmark model from [15] outperforms the *T2D* strategy in all DM A-E. Lastly, it is important to acknowledge that distinct risk profiles of DM can result in different performances. Consequently, an area of future research lies in integrating DM preferences into the cost-loss matrix and corresponding performance metric. This would enable direct comparisons between decision paradigms (e.g., risk-threshold vs trade-off) or DM.

The performance of the proposed *T2D* strategy for different flexibility price increment percentages as a function of the notification time. The previous results assumed increments of 40% and 90%. The evaluation was conducted through a sensitivity analysis in Table X. In all cases, the *T2D* strategy consistently outperformed the *DN* strategy, showing an average improvement of 20.75%. Furthermore, the minor variations observed in the performance indicator γ indicate the robustness of the *T2D* strategy to changes in flexibility prices with the notification time.

TABLE X: Sensitivity analysis for different flexibility price increments (in %) as a function of the notification time

Strategy	booking 12 hours later	2 hours before delivery	γ [M€]
DN	-	-	5.048
T2D	40%	90%	3.984
T2D	40%	110%	4.000
T2D	40%	70%	3.979
T2D	20%	70%	4.015
T2D	70%	110%	4.023

VI. CONCLUSIONS

This paper describes a predictive methodology for flexibility procurement to manage technical constraints under forecast uncertainty and consider the different flexible resources. The main contribution is a complete methodology to guide the human operator along a) the flexibility options available in each hour, ranking them according to their effectiveness in an uncertain context, and b) multiple forecast updates. This methodology was tested in a public MV network, and the results showed that: a) a methodology based on uncertainty forecasts can lead to cost savings when managing grid technical constraints, and b) choosing the best moment to reserve flexibility also leads to cost savings in comparison to a strategy that reserves flexibility when it is cheaper. It is essential to note that short-term market-based flexibility procurement still has some challenges at the distribution grid level, namely: a) dealing with the radial topology in MV and LV grids, and geographical restrictions, b) lack of a mature regulatory framework for new DSO roles and coordination across different

markets, and c) upward/downward flexibility price asymmetry. Moreover, the use of DSO’s resources (e.g., OLTC, network topology reconfiguration) should be coordinated with local flexibility markets operation, notably in areas with limited (radial) geographical coverage, fewer FSPs, and lower liquidity. Uncoordinated use of such resources may deter participation, or lead to bid and/or settlement distortions.

Topics for future work are: include the look-ahead impact (i.e., $> t + k$ of activating flexibility in lead-time $t + k$ in the flexibility options ranking; develop new metrics to evaluate decision quality under uncertainty since, in this work, the used cost-loss matrix does not integrate the operator attitude towards risk; integrate low-probability and high-impact weather-related events forecasted with NWP ensembles and develop appropriate decision-aid strategies tailored for managing such events. Moreover, the original concept of *second-level forecasting* has room for further improvement by exploring other statistical learning models and additional features; it can also be applied to different use cases like deciding to offer all RES in the day-ahead market or wait for the intraday sessions.

APPENDIX A FLEXIBILITY COST

The cost of system operator flexibility assets is computed per switching action as follows [38]:

$$C_{s/b} = \frac{1}{T_T} (F_{s/b} + \frac{F_{OT} \cdot a'_T}{t_{OT}}) \quad (17)$$

where T_T , a'_T and t_{OT} are the total allowable adjustment times, lifetime after step changed T_T times, and maintenance period, respectively. The maintenance cost and capital cost of the element (e.g., OLTC, power line, capacitor bank) is denoted by F_{OT} and $F_{s/b}$. Note that for the OLTC, the $F_{s/b}$ is equal to $F_{T_{OLTC}} \cdot (a_T - a'_T) / a_T$, where a_T is lifetime when the tap is never adjusted and $F_{T_{OLTC}}$ is the capital cost of the transformer with OLTC. Table XI presents the parameters of the flexibility cost formula per each element.

TABLE XI: Parameters for calculating flexibility cost of SO assets

Parameter	Shunt (CB/RS)	OLTC	Power lines
a_T (year)	-	45	-
a'_T (year)	20	35	20
T_T (times)	14600	89425	7300
t_{OT} (year)	4	17.5	2
F_{OT} (€/times)	600	300	60
$F_{s/b}$ (€/MVar)	7500	11250	15300

REFERENCES

- [1] H. Holttinen, A. Groom, E. Kennedy, D. Woodfin, L. Barroso, A. Orths, and et al., “Variable renewable energy integration: Status around the world,” *IEEE Pow. and Ener. Mag.*, vol. 16, no. 9, pp. 86–96, Nov. 2021.
- [2] G. Doorman, G. Thorpe, C. Dornellas, C. Correa, M. Laasri, J. Richstein, and et al., “Short-term flexibility in power systems: drivers and solutions,” *Cigre, WG C5.27, Tech. Rep. TB 808*, 2020.
- [3] F. Genoese, R. Kielak, G. Trienekens, K. Poncelet, O. A. Garcia, K. Kukkk, and et al., “Roadmap on the evolution of the regulatory framework for distributed flexibility,” *ENTSO-E, CEDEC, E.DSO, Eurelectric, GEODE, Tech. Rep.*, 2021.

- [4] O. Valarezo, T. Gómez, J. P. Chaves-Avila, L. Lind, M. Correa, D. U. Ziegler, and R. Escobar, "Analysis of new flexibility market models in Europe," *Energies*, vol. 14, no. 12, p. 3521, 2021.
- [5] A. Marot, A. Kelly, M. Naglic, V. Barbesant, J. Cremer, A. Stefanov, and J. Viebahn, "Perspectives on future power system control centers for energy transition," *J. of Mod. Pow. Sys. and Clean Ener.*, vol. 10, no. 2, pp. 328–344, 2022.
- [6] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo, "Power systems optimization under uncertainty: A review of methods and applications," in *PSCC 2022*, Porto, Portugal, 2022.
- [7] H. Zhang and P. Li, "Chance constrained programming for optimal power flow under uncertainty," *IEEE Trans. on Power Sys.*, vol. 26, no. 4, pp. 2417–2424, Nov. 2011.
- [8] L. Roald, M. Vrakopoulou, F. Oldewurtel, and G. Andersson, "Risk-constrained optimal power flow with probabilistic guarantees," in *PSCC 2014*, Wroclaw, Poland, Aug. 2014.
- [9] L. Roald and G. Andersson, "Chance-constrained AC optimal power flow: Reformulations and efficient algorithms," *IEEE Trans. on Power Sys.*, vol. 33, no. 3, pp. 2906–2918, May 2018.
- [10] I. Mezghani, S. Misra, and D. Deka, "Stochastic AC optimal power flow: A data-driven approach," *Elect. Pow. Sys. Res.*, vol. 189, p. 106567, Dec. 2020.
- [11] T. Soares, R. J. Bessa, P. Pinson, and H. Morais, "Active distribution grid management based on robust AC optimal power flow," *IEEE Trans. on Smart Grid*, vol. 9, no. 6, pp. 6229–6241, Nov. 2018.
- [12] Y. Guo, K. Baker, E. Dall'Anese, Z. Hu, and T. H. Summers, "Data-based distributionally robust stochastic optimal power flow – Part I: Methodologies," *IEEE Trans. on Power Sys.*, vol. 34, no. 2, pp. 1483–1492, Mar. 2019.
- [13] A. Marot, B. Donnot, G. Dulac-Arnold, A. Kelly, A. O'Sullivan, J. Viebahn, M. Awad, I. Guyon, P. Panciatici, and C. Romero, "Learning to run a power network challenge: a retrospective analysis," in *Proc. of Mach. Learn. Res.*, vol. 133, 2021, pp. 112–132.
- [14] M. Hasquenoph, S. Hourrig, and O. Carré, "Active power curtailment for MV network operational planning in an industrial environment," *CIRED - Open Access Proc. J.*, vol. 1, pp. 1554–1557, Oct. 2017.
- [15] E. Memmel, S. Schluters, R. Volker, F. Schuldt, K. V. Maydell, and C. Agert, "Forecast of renewable curtailment in distribution grids considering uncertainties," *IEEE Access*, vol. 9, pp. 60 828–60 840, 2021.
- [16] A. H. Murphy and Q. Ye, "Optimal decision making and the value of information in a time-dependent version of the cost-loss ratio situation," *Monthly Weather Rev.*, vol. 118, no. 4, pp. 939–949, Apr. 1990.
- [17] C. Wanke and D. Greenbaum, "Sequential congestion management with weather forecast uncertainty," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, Jun 2008.
- [18] S. Jewson, S. Scher, and G. Messori, "Decide now or wait for the next forecast? testing a decision framework using real forecasts and observations," *Mont. Weat. Rev.*, Mar. 2021.
- [19] —, "Communicating properties of changes in lagged weather forecasts," *Weath. and Forec.*, vol. 37, no. 1, pp. 125–142, Jan. 2022.
- [20] P. Tankov and L. Tinsi, "Stochastic optimization with dynamic probabilistic forecasts," *Ann. of Oper. Res.*, Aug. 2022.
- [21] J. Bellenbaum and J. C. Weber, "Designing flexibility procurement markets for congestion management – investigating two-stage procurement auctions," *Ener. Econ.*, vol. 106, p. 105775, Feb. 2022.
- [22] T. Muhlporf, V. Hagenmeyer, and T. Faulwasser, "The price of uncertainty: Chance-constrained OPF vs. in-hindsight OPF," in *PSCC 2018*, Dublin, Ireland, Jun. 2018.
- [23] S. Chondrogiannis, J. Vasiljevska, A. Marinopoulos, I. Papaioannou, and G. Flego, "Local electricity flexibility markets in Europe," Joint Research Centre, Tech. Rep. JRC130070, Oct. 2022.
- [24] I.-I. Avramidis, F. Capitanescu, and G. Deconinck, "Thoughts on conceiving a "fair" value for flexibility," *Int. J. of Elect. Pow. & Ener. Sys.*, vol. 147, p. 108893, May 2023.
- [25] Y. Ruwaida, J. P. Chaves-Avila, N. Etherden, I. Gomez-Arriola, G. Gürses-Tran, K. Kessels, and et al., "TSO-DSO-customer coordination for purchasing flexibility system services: Challenges and lessons learned from a demonstration in Sweden," *IEEE Trans. on Power Sys.*, In Press, 2022.
- [26] P. Girvan, A. Stolte, and M. Mann, "Conceptualising different approaches to the new redispatch procedure in Germany with special regard to the distribution of competences among the grid operators," in *IEEE Pow. and Ener. Student Summit*, 2020.
- [27] C.-L. Hwang and K. Yoon, *Methods for Multiple Attribute Decision Making*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 58–191.
- [28] B. Hill, "Confidence and decision," *Games and Econ. Behav.*, vol. 82, pp. 675–692, 2013.
- [29] K. Christakou, J.-Y. LeBoudec, M. Paolone, and D.-C. Tomozei, "Efficient computation of sensitivity coefficients of node voltages and line currents in unbalanced radial electrical distribution networks," *IEEE Trans. on Smart Grid*, vol. 4, no. 2, pp. 741–750, 2013.
- [30] J. Grainger and W. Stevenson, *Power System Analysis*, ser. Electrical engineering series. McGraw-Hill, 1994.
- [31] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. on Power Sys.*, vol. 24, no. 1, pp. 40–49, Feb. 2009.
- [32] J. Naughton, H. Wang, M. Cantoni, and P. Mancarella, "Co-optimizing virtual power plant services under uncertainty: A robust scheduling and receding horizon dispatch approach," *IEEE Trans. on Power Sys.*, vol. 36, no. 5, pp. 3960–3972, Sep. 2021.
- [33] J. B. and, "Market-based redispatch in the distribution grid. why it works," DNV GL Energy, Tech. Rep. 2020-0977, Sep. 2020.
- [34] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," *IEEE Trans. on Sust. Ener.*, vol. 8, no. 4, pp. 1571–1580, 2017.
- [35] C. Gonçalves, L. Cavalcante, M. Brito, R. J. Bessa, and J. Gama, "Forecasting conditional extreme quantiles for wind energy," *Elect. Pow. Sys. Res.*, vol. 190, p. 106636, Jan. 2021.
- [36] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv:1406.1078*, 2014.
- [37] M. Ni, J. D. McCalley, V. Vittal, and T. Tayyib, "Online risk-based security assessment," *IEEE Trans. on Power Sys.*, vol. 18, no. 1, pp. 258–265, Apr. 2003.
- [38] Y. jun Zhang and Z. Ren, "Optimal reactive power dispatch considering costs of adjusting the control devices," *IEEE Trans. on Power Sys.*, vol. 20, no. 3, pp. 1349–1356, Aug. 2005.
- [39] P. C. Fishburn, "Methods of estimating additive utilities," *Manag. Sci.*, vol. 13, no. 7, pp. 435–453, Mar. 1967.
- [40] Z. Wang, "Iowa distribution test systems," December 2022. [Online]. Available: <http://wzy.ece.iastate.edu/Testsystem.html>
- [41] N. Marinho and Y. Phulpin, "Economics of distributed flexibilities for the grid considering local short-term uncertainties," in *2015 IEEE Eindhoven PowerTech*, Eindhoven, Netherlands, 2015.
- [42] Y. Zhu, Z. Toth, R. Wobus, D. Richardson, and K. Mylne, "The economic value of ensemble-based weather forecasts," *Bull. of the Amer. Meteor. Soc.*, vol. 83, no. 1, pp. 73–84, Jul. 2002.
- [43] A. Longo, T. Bouman, S. Giaccaria, and T. Efthimiadis, "Societal appreciation of energy security: Volume 1: Value of lost load ? residential consumers (EE, NL and PT)," European Commission, Luxembourg (Luxembourg), Tech. Rep., 2018.
- [44] P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, no. 2, pp. 119–132, Apr. 2004.

Ricardo J. Bessa (M'18–SM'19) received the *Licenciado* degree in Electrical and Computer Engineering, the M.Sc. degree in Data Analysis and Decision Support Systems, and the Ph.D. degree in Sustainable Energy Systems from the University of Porto. He is the coordinator of the Center for Power and Energy Systems, INESC TEC. His research interests include renewable energy, energy analytics, smart grids, and electricity markets. He is an Associate Editor for the *IEEE Transactions on Sustainable Energy*.

Ferinar Moaidi received the B.Sc. and M.Sc. degree in Electrical Engineering-Power Systems from K. N. Toosi University of Technology, Iran, in 2016 and 2018. She is pursuing a Ph.D. in Electrical and Computer Engineering at the University of Porto and is a Researcher with the Center for Power and Energy Systems, INESC TEC. Her research interests include power systems, smart grids, electricity markets, and renewable energy.

João Viana received a B.Sc degree in Mathematics from the University of Minho, in 2008, and a M.Sc. degree in Computer Science from the Faculty of Sciences of the University of Porto, in 2019. Currently, he is a Researcher with the Center for Power and Energy Systems, INESC TEC. His research interests include data engineering and modeling and energy analytics.

José R. Andrade received the M.Sc. degree in Electrical and Computer Engineering from the Faculty of Engineering of the University of Porto, Portugal, in 2016. He is a Researcher with the Center for Power and Energy Systems, INESC TEC. His research interests include energy analytics, renewable energy, electricity prices forecasting, and the potential applications of natural language processing to the energy sector.