Very-Short-Term Probabilistic Forecasting for a Risk-Aware Participation in the Single Price Imbalance Settlement

Jérémie Bottieau, *Student Member, IEEE*, Louis Hubert, Zacharie De Grève, *Member, IEEE*, François Vallée, *Member, IEEE*, and Jean-François Toubeau, *Member, IEEE*

***Abstract*—The single imbalance pricing is an emerging mechanism in European electricity markets where all positive and negative imbalances are settled at a unique price. This real-time scheme thereby stimulates market participants to deviate from their schedule to restore the power system balance. However, exploiting this market opportunity is very risky due to the extreme volatility of the real-time power system conditions. In order to address this issue, we implement a new tailored deep-learning model, named encoder-decoder, to generate improved probabilistic forecasts of the imbalance signal, by efficiently capturing its complex spatio-temporal dynamics. The predicted distributions are then used to quantify and optimize the risk associated with the real-time participation of market players, acting as price-makers, in the imbalance settlement. This leads to an integrated forecast-driven strategy, modeled as a robust bi-level optimization. Results show that our probabilistic forecaster achieves better performance than other state of the art tools, and that the subsequent risk-aware robust dispatch tool allows finding a tradeoff between conservative and risk-seeking policies, leading to improved economic benefits. Moreover, we show that the model is computationally efficient and can thus be incorporated in the very-short-term dispatch of market players with flexible resources.**

***Index Terms*—Deep learning, electricity markets, encoder-decoder, robust optimization, single imbalance pricing.**

NOMENCLATURE

A. Sets and Indices

R^+	Set of up-regulation blocks offered in the imbalance settlement, index r^+
R^-	Set of down-regulation blocks offered in the imbalance settlement, index r^-
Υ	Set of flexible units, index υ

B. Upper-level Decision Variables: Θ_U

$\mathbf{e}^{imb,+}, \mathbf{e}^{imb,-}$	Positive (negative) imbalance energy of the market participant, [MWh]
$\Delta p_\upsilon^+, \Delta p_\upsilon^-$	Real-time upward (downward) power deviation of unit υ, [MW]

C. Lower-level Decision Variables: Primal Θ_{L1} and Dual Θ_{L2}

s_{r^+}, s_{r^-}	Dispatch of upward (downward) regulation power, pertaining to block r^+ (r^-), [MWh]
λ^{SI}	Dual variable reflecting the market clearing imbalance price, [€ /MWh]
μ_{r^+}, μ_{r^-}	Dual variable reflecting the economic surplus associated with block r^+ (r^-), [€ /MWh]

D. Parameters

Δt	Time resolution, [15 minutes]
$\Lambda_{r^+}, \Lambda_{r^-}$	Activation price corresponding to up (down) regulation block r^+ (r^-), [€ /MWh]
S_{r^+}, S_{r^-}	Upward (downward) regulation limit for block r^+ (r^-), [MWh]
$\hat{SI}$	Predicted system imbalance, decomposed into its positive $\widehat{SI}^+$ and negative $\widehat{SI}^-$ components, [MWh]
$C_\upsilon^+, C_\upsilon^-$	Up (down) costs of unit υ, [€ /MWh]
$\Delta P_\upsilon^P, \Delta P_\upsilon^E, \Delta P_\upsilon^R$	Limits in terms of power *P*, energy *E* and ramping ability *R* of unit υ, converted into the maximum deviation of the power level at each time step, [MW]

Manuscript received January 9, 2019; revised May 15, 2019 and July 18, 2019; accepted September 5, 2019. Date of publication September 11, 2019; date of current version February 26, 2020. This work was supported by the Energy Transition Funds under Project EPOC 2030-2050 organized by the FPS economy, S.M.E.s, Self-employed and Energy. Paper no. TPWRS-00037-2019. *(Corresponding author: Jérémie Bottieau.)*

The authors are with the Electrical Power Engineering Unit, University of Mons, 7000 Mons, Belgium (e-mail: jeremie.bottieau@umons.ac.be; louis.hubert@student.umons.ac.be; zacharie.degreve@umons.ac.be; vallee@umons.ac.be; jean-francois.toubeau@umons.ac.be).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2019.2940756

I. INTRODUCTION

DECARBONISATION of the electricity sector is an essential milestone to respond to the climate change threat. Henceforth, policy makers are setting up ambitious targets for integrating large shares of renewables, such as solar and wind energies, into the electricity generation mix. These resources are weather-dependent, and thus uncertain and intermittent, which results in an increased need of flexibility to alleviate imbalances between load and generation in power systems [1].

In this context, properly gathering and valorizing the flexibility available within all layers of the grid is becoming crucial in order to ensure the transition towards a low-carbon system at a reasonable cost [2]. Up until now, much effort has been made to improve the optimal allocation of flexibility on a medium-term to day-ahead horizon, with a joint participation in both (competitive) energy and reserve markets [3]–[8].

However, as the electricity exchanges are becoming more and more short-term oriented due to increased uncertainties, adjustment strategies near to real-time are growing in significance. Traditionally, European market rules are such that they encourage market players to adopt a risk-averse policy, where their only objective is to exploit their real-time flexible resources to alleviate their own residual imbalances [9]–[13]. By over-incentivizing market parties to be in balance, such an approach prevents them to share the cost-savings with the market in case of helpful imbalances (that restore the energy balance within the power system). To address this issue, the single imbalance pricing is currently emerging as a more efficient mechanism [14]. In this way, the European Agency for the Cooperation of Energy Regulators, which coordinates national energy regulators and works towards the harmonization of market rules at the European level, has explicitly opted for such a single pricing as the preferred approach to monetize imbalances [15]. This mechanism has indeed proved its superior efficiency in countries where it has been adopted (e.g., Belgium, Germany) [16]. In this specific design, all imbalance positions are settled at a single price, thus stimulating parties to deviate from their schedule (to counterbalance the global grid imbalance) in order to help the system operator in maintaining the equality between the total generation and consumption in its control area (see Section II).

Practically, the participation of market players in this single pricing settlement is very challenging due to two obstacles:

i) The risk associated with the extreme volatility and unpredictability of the system imbalance. Indeed, in case of a misestimation of the future system state, the actor can adopt a position that aggravates the imbalance of the power system, thereby suffering important financial penalties. Hence, a market player needs to accurately anticipate the future aggregated position of all its rivals, while considering its own impact on the system (i.e., price-maker assumption).
ii) The time constraint when solving this near real-time optimization procedure. In this way, in order to benefit from the latest updated value of the system imbalance given by the system operator, the optimization should start no later than the beginning of the imbalance settlement period (which has a 15 minutes time span). The computation time for the decision-making process must thus not exceed 5 minutes, to give time to fast flexible resources to adjust their output power.

Because of these challenges, as well as the current sparse adoption of the single imbalance pricing, the literature is still at a very primitive stage. In [17], the participation of a wind power producer in the day-ahead market considering its price-maker contribution in the imbalance settlement is formulated as a mathematical program with equilibrium constraints. This formulation is extended in [18] where the wind player participates and behaves as a price-maker in both the day-ahead and balancing markets. However, the uncertainty is modeled through scenarios, so that the computational burden can be an issue for operational decisions. Moreover, although this approach allows to make informed day-ahead decisions in view of possible forecast errors, they do not define the player's actual real-time operation. In contrast, risk-neutral approaches are developed in [19], [20] to enable an aggregator to participate in the balancing market, showing the economic interest of such real-time strategies in a one-pricing scheme.

Overall, none of these pioneering works attempts to operate with a risk-aware perspective in the single-price imbalance settlement. In this paper, we therefore propose a forecast-driven tool, in which state-of-the-art probabilistic forecasts of the future system imbalance are integrated within the very-short-term dispatch of market actors.

In recent years, an impressive amount of work has been done on probabilistic forecasting methods for better informing the decision-making tools related to power systems and electricity markets [21]–[24]. Two distinct philosophies can be found [25]: two-step procedures based on the addition of a probabilistic forecast error to a deterministic forecasting model [26]–[29], and methods directly providing probabilistic predictions of the variable of interest [30]–[32]. In the latter approach, in a non-parametric setting, two methods have been widely applied in the power systems literature, namely kernel density estimators [33], [34] and direct quantile regression models [35], [36].

In this paper, a tailored quantile regression tool is firstly developed, which is one of the original contributions of the work. This model generates improved probabilistic predictions that fully characterize, in a non-parametric way, the uncertainty of the future system imbalance (in MW). Traditionally, the system imbalance was assumed to be equal to load and/or wind forecasting errors [37], [38] whereas, in practice, its variability is much more complex (which is mainly due to the aggregated effect of the individual errors of all market actors, potentially acting strategically).

In a second step, the predicted quantiles are used to quantify risks and optimize the participation of a market player in the imbalance settlement. Such an integrated procedure allows to directly quantify the potential of realistic forecasts on the performance of actual optimization strategies.

In this respect, the contributions of this work are three-fold:

i) Implement a decision-making tool, based on a bi-level model, aiming at adjusting the real-time position of a market player acting as a price-maker (i.e., its own imbalance affects the overall system position, and thus the imbalance tariff). The procedure is embedded within a robust optimization framework, to ensure that the solution is resilient to forecast uncertainty. Indeed, such an approach is intrinsically designed to hedge against the risk associated with extreme scenarios (arising from the high variability of system positions). The resulting model is a mixed-integer linear program (MILP) that is computationally efficient (compact) in comparison with alternatives such as scenario-based optimization, to

enable the practical utilization of the tool within very short timeframes.

ii) Generate reliable predictions of the future system imbalance, under the form of probabilistic intervals, to properly guide the robust optimization tool. Specifically, the goal is to implement an improved forecaster, named encoder-decoder, which is built upon our previous work [39] and recent breakthroughs in sequence-to-sequence recurrent neural networks, with the goal of optimally extracting the complex dynamic characteristics of the system imbalance. Forecasting this variable is a new and challenging task, and our results show that the proposed tool provides accurate forecasts (with sharp intervals) in comparison with 8 state-of-the-art techniques such as ensemble methods (XGBoost and quantile regression forests) and deep neural networks.

iii) Offer a probabilistic guarantee for the robust solution, using the predicted quantiles of the system imbalance. Indeed, these quantiles enable the market player to choose the level of risk during the optimization, thereby providing a user-defined trade-off between conservatism and economic efficiency [40].

The proposed integrated approach combines therefore the strengths of probabilistic forecasting and robust optimization to form an efficient, fast, risk-aware decision procedure. Outcomes from a numerical out-of-sample analysis show that the proposed forecast-driven methodology allows to increase the benefits of market players with flexible resources available in real-time, thereby enhancing the value of flexibility. Such improved dispatch procedures may thus help to valorize new flexible resources such as load-shifting capabilities.

The rest of the paper is organized as follows. Section II introduces the market environment and motivates the proposed approach. Section III firstly describes the new deep learning based tool to generate very short-term probabilistic predictions of the system imbalance signal. Then, we present the bi-level robust optimization framework for a risk-aware participation in the single price imbalance settlement. The results in terms of both prediction accuracy and quality of operational decisions are discussed in Section IV. Finally, conclusions and perspectives for future research are provided in Section V.

II. Motivation

Following the liberalization of the European electricity sector, the task of balancing electricity generation and consumption (to guarantee the stability of the power system) is supported by market actors, which are named Balance Responsible Parties (BRPs). To help them in achieving this objective, the European market architecture is organized into separate energy-only and operating reserves services, which are traded sequentially via independent auctions (Fig. 1).

In this system, BRPs are responsible to continuously equalize their own load and generation, taking into account the electricity traded with other BRPs in the different energy-only market floors (i.e., long-term, day-ahead and intraday stages). Then, the real-time residual imbalance at the system level is compensated through a balancing mechanism, in which the system operator uses power reserves made available by some BRPs (box A. in Fig. 1) through the reserve capacity market.

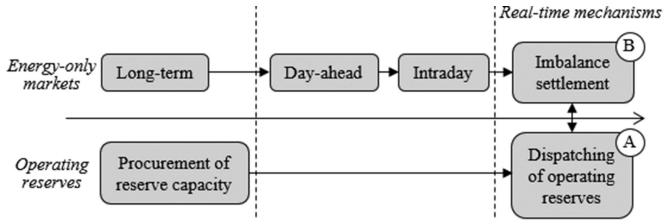

Fig. 1. Structure of the considered European market design.

TABLE I
Single Price Imbalance Settlement

P = Production C = Consumption SI = System Imbalance		Transmission System Operator (TSO) control area	
		Excess (P>C) Positive SI	Shortage (P<C) Negative SI
BRP area	Excess	MDP (low price) TSO pays BRP	MIP (high price) TSO pays BRP
	Shortage	MDP (low price) BRP pays TSO	MIP (high price) BRP pays TSO

The costs incurred by the real-time activation of reserves are covered by applying a charge to each unbalanced BRP. These charges, computed (averaged) on a quarter-hourly basis, are applied as soon as an imbalance occurs (box B. in Fig. 1). Therefore, this mechanism is not a market *per se* (with bids/offers in a competitive environment), and is referred to as imbalance settlement. European regulatory bodies have recently defined guidelines to penalize such imbalances using a single pricing settlement [15]. In this paper, we thereby consider this single price imbalance settlement design, where all imbalance positions are settled at a unique price. This system favors BRPs that help at restoring the imbalance of the system (green boxes in Table I), while penalizing those who are aggravating the imbalance situation (red boxes in Table I).

In general, the (single) imbalance price λ^{SI} depends on the actual amount of reserves activated by the system operator. When there is a generation shortage in the grid, the system operator must activate upward reserves, the cost of which is defined as the marginal incremental price (MIP). The BRPs which are responsible of this shortage must pay this MIP (which is typically higher than the day-ahead market price, and thus less economic), whereas actors with an excess of generation receive this (very attractive) MIP. Similarly, when there is a generation surplus at the system level, the subsequent activation of downward reserve results in the marginal decremental price (MDP). The BRPs in generation surplus receive this MDP (which is typically lower, and thus, less profitable than the day-ahead energy price) for their surplus, whereas BRPs in generation shortage will pay this (attractive) small MDP fee for their corresponding negative imbalance, since they help at restoring the system balance.

The imbalance price λ^{SI} is thus subject to abrupt regime switching between MDP and MIP (as further discussed in

Fig. 3 in Section III.B), which makes its prediction very unstable and difficult. This work thereby focuses on the imbalance volume prediction (in MW) in Section III.A, which is subsequently interlinked with its corresponding price within the robust optimization procedure in Section III.B.

III. Model Description

A. Probabilistic Forecasts

The objective of this part is to generate predictions of the system imbalance (residual imbalance of all market participants) in order to help a market player guiding its robust optimal dispatch strategy. In [41], a methodology which combines classical and data mining techniques is used to provide deterministic forecasts of the system imbalance volumes. This work puts into light the vital need of improved forecasting tools to better manage the high uncertainty of the imbalance signal. Here, probabilistic predictions are developed to endogenously capture and characterize the uncertainty associated with point predictions, originating from both the noise in explanatory variables (e.g., due to the chaotic nature of weather conditions) and the model misspecifications. In this way, we want to solve the following time series probabilistic regression problem:

$$p\left(y_{t:}|\, y_{:t}, x^{(h)}_{:t,i}, x^{(s)}_{i}, x^{(f)}_{t:,i}\right) \tag{1}$$

where $y_{t:}$ is the system imbalance to forecast (and y_t its past observed values), $x^{(h)}_{:t,i}$ are the past values of explanatory variables (e.g., recent forecast errors of renewable generation), $x^{(s)}_{i}$ are the static time-invariant features (e.g., the location of renewable sources), and $x^{(f)}_{t:,i}$ is the known future information. This knowledge on the future contains both seasonal features (e.g., the hour of the day, the day of the week, or holidays), and known events (e.g., commitment/schedules of power plants).

Generally, four types of forecasting models are proposed in the literature [42]: (i) persistence methods, which simply consider that future observations will have the same values as the current instance, (ii) physical methods, which are based on a detailed mathematical description of the environment governing the variable of interest (e.g., full modeling of the market rules and participant's behaviors to predict market prices), (iii) statistical methods, which build mathematical models of predefined complexity based on the inference (from historical observations) of basic statistics such as autocorrelation, and (iv) machine learning methods, which are based on generic nonlinear models that are trained using historical data through a self-learning procedure (without being explicitly programmed to achieve the prediction task, with no arbitrary assumptions on the model complexity).

Overall, persistence methods are very naïve and do not provide useful information for decision-making. Physical models, on the other hand, are characterized by a high computational complexity, which hinders their practical utilization for very-short-term predictions. Moreover, they require to explicitly model the behavior of all involved market parties, which cannot be easily estimated. Statistical models are usually much simpler linear forecasters, but this prevents them from capturing nonlinear dependencies and from properly representing high frequency effects (such as fast ramping events). For these reasons, machine learning approaches have recently achieved much better results than the other models, which was further fostered by the increase of reliable databases and of computer capabilities.

This work focuses on neural networks (although the case study in Section IV.A integrates and compares other state-of-the-art approaches). Indeed, in addition to their theoretical ability to capture and represent hidden mechanisms of any complexity between inputs-outputs, neural networks are flexible tools that can be tailored to the specificities of the problem, thereby improving their performance [43]. This property has seen the emergence of recurrent neural networks, advanced deep learning structures that are characterized by architectural features specifically designed to propagate relevant information from past inputs (through a powerful dynamical memory). Such recurrent networks have shown a high potential in processing and predicting complex times series data with multi-scale dynamics (akin to the system imbalance). This type of neural networks relies on a hidden (memory) state h, which captures and propagates the dynamics of the variable over the time horizon T. In this way, at each time step t, given an input sequence $\mathbf{x} = (x_1, \ldots, x_{\mathrm{T}})$ of length T, the hidden state h_t is updated by:

$$h_t = f\left(h_{t-1}, x_t\right) \tag{2}$$

where f is a nonlinear activation function (e.g., a logistic sigmoid function). It should be noted that such models can be efficiently trained using parallel calculation. Moreover, once the learning phase is achieved, the computational burden to yield real-time predictions can be neglected ($<$1 second).

However, recurrent networks are known to struggle in accessing time dependencies more than a few time steps long. The problem, commonly known as the vanishing gradient problem, arises from the fact that back-propagated errors during the training stage either fades or blows up over time. Two state-of-the-art recurrent networks, i.e., Long Short-Term Memory (LSTM) and gated recurrent units (GRU) networks, tackle this problem by using internal memory cells, whose content is controlled by gating units, thereby optimally regulating the flow of information that is propagated through time [44], [45]. Both variants rely on the same principle, but GRU networks have a simpler structure for equivalent performance levels [46].

Despite its superior accuracy in the current literature [47], the GRU architecture still suffers from two main drawbacks. Firstly, it intrinsically requires time sequential data as inputs, whereas some explanatory variables are time-invariant by nature, which prevents the GRU from fully exploiting this valuable information. Secondly, GRU networks are designed to process a fixed-size input vector for each time step, and it is thus very complex and inefficient to encode both past and future information into the model (without resorting to coding tricks that degrade the predictive accuracy).

Both these issues can be efficiently handled using an innovative architecture, referred to as encoder-decoder [48]. As represented in Fig. 2, this tool is composed of two different networks, and has recently shown promising results for translation tasks and speech recognition applications [48].

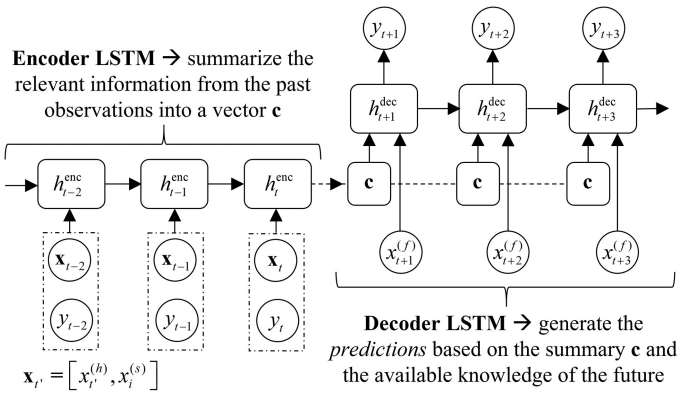

Fig. 2. Encoder-decoder architecture to provide time series prediction.

The temporal nature of the encoder is used to process the past observed values y_t and $x^{(h)}_{:t,i}$, as well as the horizon-agnostic context $x^{(s)}_i$ in order to extract the relevant historical information. Practically, this information is contained into a reduced vector $\mathbf{c}$ of fixed dimensions, based on the last hidden state h^{enc}_t. Then, the decoder leverages this representation $\mathbf{c}$, along with the known future $x^{(f)}_{t:,i}$ information, to generate K (multi-step ahead) predictions. Overall, the hidden state of the decoder and the output sequence at time t are computed by:

$$h^{\text{dec}}_{t+k} = f_{GRU}\left(h^{\text{dec}}_{t+k-1}, x^{(f)}_{t+k,i}, \mathbf{c}\right) \tag{3}$$

$$\hat{\mathbf{y}}_{t+k} = g\left(h^{\text{dec}}_{t+k}\right) \tag{4}$$

where K is the length of the multi-horizon forecast (i.e., the number of look-ahead times k), f_{GRU} is the nonlinear function representing the GRU neural network, and g is a linear transformation of the GRU outputs h^{dec}_{t+k}.

Then, the model has to be adjusted to obtain distributions characterizing the prediction uncertainty. In this work, we use quantile regression to predict the specified quantiles $q \in Q$ of the target distribution, i.e.,:

$$q = \mathbb{P}\left(y_{t+k} \leq \hat{y}^{(q)}_{t+k} \,\middle|\, y_{:t}, x^{(h)}_{:t,i}, x^{(s)}_i, x^{(f)}_{t+k,i}\right), \quad \forall k \in K \tag{5}$$

Practically, the model is trained with the historical database to minimize the quantile loss that provides the values $\hat{y}^{(q)}$ corresponding to the different quantiles q of interest. However, in case of perfect prediction, the quantile loss cannot be differentiated. To address this issue, we construct a smooth approximation of the pinball loss [49], by including the Huber norm within the loss function [50]. The idea is to replace the L1 norm by the (continuously differentiable) L2 norm when the error is lower than a (preferably small) user-defined threshold ε (in this paper, we arbitrarily use $\varepsilon = 10^{-6}$):

$$H(y_k, \hat{y}^{(q)}_k) = \begin{cases} \frac{\left(\hat{y}^{(q)}_k - y_k\right)^2}{2\varepsilon} & \left|\hat{y}^{(q)}_k - y_k\right| \leq \varepsilon \\ \left|\hat{y}^{(q)}_k - y_k\right| - \frac{\varepsilon}{2} & \left|\hat{y}^{(q)}_k - y_k\right| > \varepsilon \end{cases} \tag{6}$$

where the values $\hat{y}^{(q)}$ are the outputs of the forecaster, and y the actual (ground truth) observations. The approximated pinball loss E_Q can then be calculated as:

$$E_Q = \sum_{k\in K}\sum_{q\in Q} \begin{cases} q \cdot H\left(y_k, \hat{y}^{(q)}_k\right) & \hat{y}^{(q)}_k < y_k \\ (1-q) \cdot H\left(y_k, \hat{y}^{(q)}_k\right) & \hat{y}^{(q)}_k \geq y_k \end{cases} \tag{7}$$

As the resulting loss function is differentiable, the neural networks can be trained using efficient gradient-based algorithms. The conditional quantiles provided by the developed methods will give the possibility to select different levels of probabilistic guarantees as input parameters for the robust optimization framework.

B. Robust Optimization Formulation

Once the predicted distributions are obtained, the purpose is to adequately exploit them to feed the optimization tool for the strategic participation of a Balance Responsible Party (BRP, i.e., a merchant energy storage actor), in the imbalance settlement. This near-real-time stage consists in evaluating whether it is profitable for the market actor (under risk-constraints) to deviate from its balanced position by increasing or decreasing the output power of its flexible resources. It should be noted that even if a unit is operating at its maximum output power, flexibility is still available in the opposite direction (i.e., reduction of power). This situation allows, e.g., selling at a high price in the day-ahead market, while buying low in the imbalance settlement.

In order to properly consider the fact that the player can affect imbalance prices (and thus its own profit) through its contribution, the market-clearing process is explicitly included in the agent's optimization problem. The resulting approach is formulated as a bi-level model, which is characterized by an upper-level problem that maximizes the profit of the market player by identifying its optimal strategic imbalance position, and a lower-level problem that simulates the (single price) quarter-hourly settlement of energy imbalances.

Overall, the market participant optimizes its position by anticipating what could be the system imbalance (at the end of the considered quarter of an hour), given its own decisions and the (exogenous) residual deviation of all its rivals $\widehat{SI}$. In turn, the system imbalance is cleared with the knowledge of the player's contribution. This (single time-step) bi-level optimization is run sequentially (96 times a day) at the beginning of each 15 minutes imbalance settlement period, in order to use the last measured values of the system imbalance communicated by the system operator. The deterministic formulation is summarized in (8)–(9):

$$\max_{\Theta_U} \lambda^{SI}\left(e^{imb,+} - e^{imb,-}\right) - \sum_{\upsilon\in\Upsilon}\left(C^+_\upsilon \Delta p^+_\upsilon - C^-_\upsilon \Delta p^-_\upsilon\right)\Delta t \tag{8a}$$

subject to:

$$e^{imb,+} = \sum_{\upsilon\in\Upsilon} \Delta p^+_\upsilon \Delta t \leq S\hat{I}^- \tag{8b}$$

$$e^{imb,-} = \sum_{\upsilon\in\Upsilon} \Delta p^-_\upsilon \Delta t \leq S\hat{I}^+ \tag{8c}$$

$$\widehat{SI}^{+} = \widehat{SI} \text{ if } \widehat{SI} > 0, \text{ and } \widehat{SI}^{+} = 0 \text{ if } \widehat{SI} < 0$$

$$\widehat{SI}^{-} = \widehat{SI} \text{ if } \widehat{SI} < 0, \text{ and } \widehat{SI}^{-} = 0 \text{ if } \widehat{SI} > 0 \quad (8d)$$

$$0 \leq \Delta p_{\upsilon}^{+} \leq \min\left(\Delta P_{\upsilon}^{P+}, \Delta P_{\upsilon}^{E+}, \Delta P_{\upsilon}^{R+}\right) \quad \forall \upsilon \quad (8e)$$

$$0 \leq \Delta p_{\upsilon}^{-} \leq \min\left(\Delta P_{\upsilon}^{P-}, \Delta P_{\upsilon}^{E-}, \Delta P_{\upsilon}^{R-}\right) \quad \forall \upsilon \quad (8f)$$

$$\text{Technology}-\text{specific constraints} \quad (8g)$$

$$\arg\min_{\Theta_{L1}} \left\{ \sum_{r^+ \in R^+} \Lambda_{r^+} s_{r^+} - \sum_{r^- \in R^-} \Lambda_{r^-} s_{r^-} \right\} \quad (9a)$$

subject to:

$$\sum_{r^+ \in R^+} s_{r^+} - \sum_{r^- \in R^-} s_{r^-} = -\left(e^{imb,+} - e^{imb,-}\right) - \hat{SI} : \lambda^{SI} \quad (9b)$$

$$0 \geq -s_{r^+} \geq -S_{r^+} : \mu_{r^+}, \quad \forall r^+ \quad (9c)$$

$$0 \geq -s_{r^-} \geq -S_{r^-} : \mu_{r^-}, \quad \forall r^- \quad (9d)$$

with Θ_U, the set of decision variables $e^{imb,+}$, $e^{imb,-}$, Δp_{υ}^{+} and Δp_{υ}^{-} of the upper-level problem, and Θ_{L1}, the decision variables s_{r+} and s_{r-} of the (primal) lower-level problem. The dual variables Θ_{L2} of the lower-level, i.e., λ^{SI}, μ_{r+} and μ_{r-}, are indicated after each constraint, preceded by a colon.

The objective function (8a) is the maximization of the revenues from the imbalance settlement (i.e., by multiplying the single imbalance price λ^{SI} with the imbalance volumes $e^{imb,+}$ and $e^{imb,-}$ of the market player), considering the intrinsic costs C_{υ}^{+} and C_{υ}^{-} of all its units. Constraints (8b)–(8c) allow to obtain a cost-optimal allocation of the available resources, by activating the most cost-effective units. For instance, extra power will be provided by units with low marginal costs, whereas power reduction will be performed by expensive units (to save their high operating costs). Together with (8d), these constraints also ensure that the actor's contribution helps at restoring the system balance. The upward (8e) and downward (8f) flexibility that can be provided by each unit υ is limited by operational constraints in terms of available output power ΔP_{υ}^{P}, energy capacity ΔP_{υ}^{E} and ramping ability ΔP_{υ}^{R}. Then, technology-specific constraints (8g), such as the minimum up/down times of conventional generators, have also to be integrated to ensure the global consistency of the formulation.

The lower-level problem (9) represents the quarter-hourly market clearing of energy imbalances. The objective (9a) reflects the costs incurred by the real-time activation of operating reserves (following the merit order in Fig. 3). The parameter Λ_{r+} is the price for activating the upward reserve S_{r+} (in case of negative system imbalance). This marginal incremental price increases with respect to the imbalance severity, which reflects the necessity for the system operator to activate more expensive reserves (far in the merit order) to fully restore the system frequency. In general, Λ_{r+} is higher than the day-ahead market price for all r^+, which stimulates market players to increase their generation (or decrease their consumption) to sell energy at a high price. Similarly, Λ_{r-} is the imbalance price, typically lower than the day-ahead price, associated with the activation of

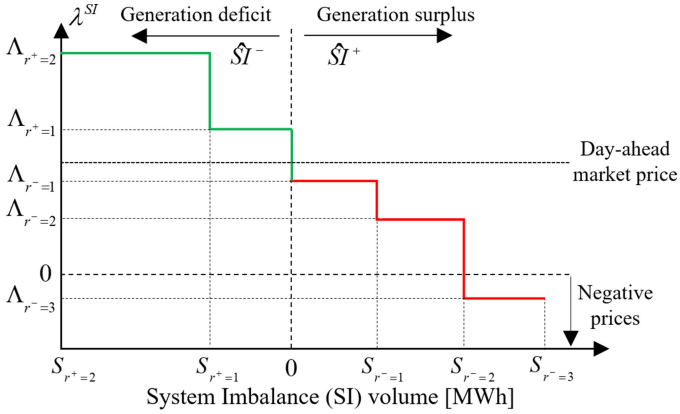

Fig. 3. Single imbalance prices per volume levels for one quarter-of-an-hour, made publicly available by the system operator.

downward reserves S_{r-} (in case of positive system imbalance). This marginal decremental price decreases (possibly down to negative values) for large surplus of generation, reflecting the temporary low value of electricity.

Following the European guidelines to promote a transparent and more competitive market, the Belgian system operator publishes, on a day-ahead basis with updates in intraday, information about the imbalance prices per volume level for each 15-min period [51]. Henceforth, the parameters Λ_{r+}, Λ_{r-}, S_{r+} and S_{r-} can be retrieved from these data, and are considered as parameters in the optimization.

Constraint (9b) guarantees that the activated amount of upward (s_{r+}) and downward (s_{r-}) reserves exactly offset the strategic imbalance volumes of the market participant ($e^{imb,+}$ and $e^{imb,-}$) and the expected aggregated deviation of all its rivals ($\widehat{SI}$). The dual variable associated with these balance constraints is the imbalance price λ^{SI}, which reflects the system cost to compensate one more MW in imbalance.

Constraints (9c) and (9d) ensure that reserves are activated in accordance with the merit order (Fig. 3), accounting for the volume limitations of the different block offers. The dual variables μ_{r+} (and μ_{r-}) associated with these constraints correspond to the economic surplus, i.e., the monetary gain of the system if the block S_{r+} (or S_{r-}) offers one more MW of reserve. They represent thus the difference between the imbalance price λ^{SI} and the price of Λ_{r+} (or Λ_{r-}). Hence, μ_{r+} (and μ_{r-}) are equal to zero if the system imbalance is such that reserves of level r^+ (or r^-) are not activated.

Overall, the formulation (7)–(8) cannot be solved directly by a linear optimization solver, due to the nested optimization, and to the bilinear terms λ^{SI} $(e^{imb,+} - e^{imb,-})$ in (7a). The first issue is addressed by replacing the lower-level problem (8) by its Karush-Kuhn-Tucker (KKT) conditions (since this lower-level is convex and the constitutive constraints satisfy some regularity conditions), resulting in a mathematical program with equilibrium constraints (MPEC). This set of KKT conditions includes complementary conditions, which implies nonlinearities. The latter can be linearized using a Big-*M* approach that relies on binary variables, which yields the mixed-integer linear conditions

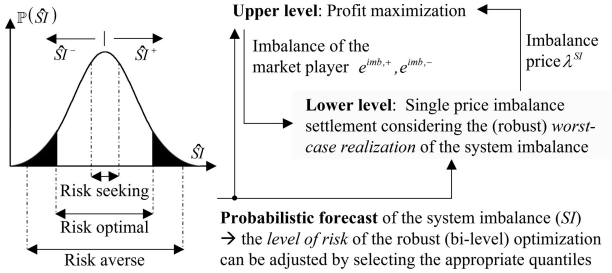

Fig. 4. Integrated (forecast-driven) bi-level robust optimization framework.

(10c)–(10l), where the *M*-parameters are large enough positive constants [17]. The selection of appropriate *M*-values can be a challenging task in the context of our optimization problem, since improper values can lead to excessively high simulation times, or even numerical ill-conditioning and violations of complementary conditions. However, practical bounds of the *M*-values can be determined based on the economical or physical interpretation of their corresponding constraints. For instance, the constraints (10h) and (10j) respectively related to the dispatch of upward s_{r+} (or downward s_{r-}) regulation power are bounded by their corresponding limits S_{r+} (or S_{r-}). Consequently, the associated *M*-values M^3_{r+} (or M^3_{r-}) can be set to these limits S_{r+} (or S_{r-}). The same logic can be applied for the activation prices in (10i) and (10k). Interestingly, in the case study presented in Section IV (run on 5472 different time steps), this approach has led to adequate results in all cases (i.e., no infeasible outcomes nor unfulfilled complementary conditions). The second source of nonlinearity, which arises from the bilinear terms in the upper-level objective function (8a), can also be expressed as a sum of linear terms (10a) by applying the strong duality equality on the market clearing problem.

In parallel, since the future system imbalance is unknown, the optimization approach must hedge against the risk of erroneous decisions that lead to severe financial penalties. This is achieved by relying on a robust formulation that immunizes the operation strategy against the worst-case realization of the system imbalance in an uncertainty set (given by the probabilistic forecasts of Section III.A). As our prediction tool provides non-parametric quantiles of the future system imbalance, the methodology allows to select a specified level of probabilistic guarantee, thereby quantifying and controlling the risk of the procedure.

Consequently, the reformulated MILP robust optimization framework (Fig. 4) can be expressed as follows:

$$\max_{\Theta} \min_{S\hat{I}} -\lambda^{SI}\widehat{SI} - \sum_{\upsilon\in\Upsilon}\left(C^+_\upsilon \Delta p^+_\upsilon - C^-_\upsilon \Delta p^-_\upsilon\right)\Delta t - \sum_{r^+\in R^+}\left(S_{r^+}\mu_{r^+} + \Lambda_{r^+}s_{r^+}\right) + \sum_{r^-\in R^-}\left(-S_{r^-}\mu_{r^-} + \Lambda_{r^-}s_{r^-}\right) \tag{10a}$$

subject to:

Market actor constraints: (8b), (8c), (8d), (8e), (8f), (8g)

$$\widehat{SI} = \widehat{SI}^+ - \widehat{SI}^- \in \left[\widehat{y}^{(q)}, \widehat{y}^{(1-q)}\right] \tag{10b}$$

$$0 \le \Lambda_{r^+} - \lambda^{SI} + \mu_{r^+} \le M^1_{r^+}(1 - z^1_{r^+}), \quad \forall r^+ \tag{10c}$$

$$0 \le s_{r^+} \le M^2_{r^+} z^1_{r^+}, \quad \forall r^+ \tag{10d}$$

$$0 \le -\Lambda_{r^-} + \lambda^{SI} + \mu_{r^-} \le M^1_{r^-}(1 - z^1_{r^-}), \quad \forall r^- \tag{10e}$$

$$0 \le s_{r^-} \le M^2_{r^-} z^1_{r^-}, \quad \forall r^- \tag{10f}$$

$$\sum_{r^+\in R^+} s_{r^+} - \sum_{r^-\in R^-} s_{r^-} = -\left(e^{imb,+} - e^{imb,-}\right) - S\hat{I} \tag{10g}$$

$$0 \le S_{r^+} - s_{r^+} \le M^3_{r^+} z^2_{r^+}, \quad \forall r^+ \tag{10h}$$

$$0 \le \mu_{r^+} \le M^4_{r^+}(1 - z^2_{r^+}), \quad \forall r^+ \tag{10i}$$

$$0 \le S_{r^-} - s_{r^-} \le M^3_{r^-} z^2_{r^-}, \quad \forall r^- \tag{10j}$$

$$0 \le \mu_{r^-} \le M^4_{r^-}(1 - z^2_{r^-}), \quad \forall r^- \tag{10k}$$

$$z^1_{r^+}, z^1_{r^-}, z^2_{r^+}, z^2_{r^-} \in \{0, 1\} \tag{10l}$$

with Θ, the set of all decision variables Θ_U, Θ_{L1} and Θ_{L2} of (8)–(9), to which are added the binary variables from (10l), needed for linearizing the KKT-related complementary conditions.

It can be shown that, for linear programming with a polyhedral uncertainty set, the optimal solution is located at one of the vertices of the uncertainty set [52]. The max-min formulation (10a) can thus be solved by an off-the-shelf optimization software that enumerates all the vertices defined by probabilistic quantiles $\hat{y}^{(q)}$ and $\hat{y}^{(1-q)}$, which allows to determine the risk-aware dispatch Δp^+_υ and Δp^-_υ.

IV. Case Study

Firstly, the quality of probabilistic forecasts (of the system imbalance) is estimated using different evaluation criteria, and a representative benchmark is built to compare the performances of both classical and state-of-the-art methods (Section IV.A). Once these probabilistic forecasts are obtained, the appropriate quantiles are integrated into the very-short-term dispatch procedure (Section IV.B). In this work, different bounds (each one being associated to a different risk-attitude) are tested to determine the optimal risk policy. The integrated (forecast & optimization) procedure is performed on actual data (2014–2018), obtained from Elia, the Belgian transmission system operator [53]. To favor reproducibility of the results, a pre-processed (cleaned) database, i.e., historical data of the system imbalance along with all its available explanatory variables used in this paper, is provided in [54].

A. Numerical Results of Probabilistic Forecasts

This subsection intends to evaluate the performance of the proposed probabilistic forecasting tool, and to compare it to other state-of-the-art techniques. All predictions (performed at the beginning of every quarter of an hour) focus on the Belgian system imbalance. Practically, a benchmark using Python 3.6.0 and the Keras library (along with the TensorFlow backend) has been implemented for the neural networks, whereas the scikit-learn library has been employed for ensemble models. The following 8 models have been tested:

- Probabilistic persistence method, where the forecast error is assumed to be random and normally distributed. Both

the mean and variance are adjusted with the last available measurements of the system imbalance.
- Auto-Regressive Integrated Moving Average (ARIMA) model (with confidence intervals), which assumes a constant variance, while time correlation is linearly represented.
- Quantile regression forest (QRF), an ensemble method that generalizes random forests[1] for estimating quantiles instead of the conditional mean [55], is used with 100 trees.
- eXtreme Gradient Boosting (XGBoost), a (multi-stage) decision tree ensemble method in which new models are sequentially created to forecast the residuals (i.e., errors) of the models optimized at the previous stage. At each stage, all models are trained (updated) together (using a gradient descent algorithm) to make the final prediction [56].
- Shallow Multi-Layer Perceptron (MLP), the traditional architecture of feedforward neural networks with 1 hidden layer, containing 450 neurons with rectifier linear units (ReLUs) as activation function.
- 3-MLP, a MLP with 3 hidden layers containing respectively 220, 300, 90 ReLUs neurons.
- Hybrid model combining a MLP with a GRU model to better capture time dependencies, with 7 cells in the recursive layer and 10 ReLUs units in the MLP layer.
- Encoder-decoder, both GRU networks with 1 hidden layer containing 12 cells.

For machine learning models, a sensitivity study (hyper-parameter optimization) was performed on several input configurations (to choose the most relevant set of explanatory variables among all publicly available information provided by the system operator) and specific parameters (to select the optimal complexity of the model) in order to ensure that the predictive potential of the different models is fully exploited. To that end, we have employed the tree-structured Parzen estimator, a sequential model-based optimization algorithm [57]. For each of the studied architectures, the neural networks are trained using the Adam algorithm [58], a stochastic gradient descent method that relies on adaptive learning rates for escaping local optima within the learning phase. In addition, all models are stabilized with regularization techniques (e.g., early stopping), to avoid overfitting.

Concerning the input configuration, a sensitivity study concludes that the past 12 historical observations of both the system imbalance y_t and historical explanatory variables $x^{(h)}_{:t,i}$ are the most suited to capture recent dynamics for this particular task. In particular, the most relevant ones are the past measured values of the total grid load, the aggregated power generation in Belgium, as well as its individual components, i.e., wind, photovoltaic (PV), nuclear and gas generation. In addition, relevant future information $x^{(f)}_{t:,i}$ is also included into the model, i.e., the same electrical quantities (grid load, wind, PV, nuclear and gas generation) forecasted (and published in intraday) by the system operator. Moreover, the models also integrate spot prices (cleared in the day-ahead market) and temporal aspects (quarter-of-an-hour of the day and day of the week). This calendar information is introduced by a mutually exclusive binary representation. For instance, 96 binary inputs are used to represent all daily quarters-of-an-hour. With such a representation, when one input is equal to 1, all others are set to 0.

All models are trained to capture the system conditions hidden in the explanatory variables, and to generate the appropriate forecast accordingly. There is therefore no need of systematically re-training the models on a short time-frame to continuously adapt to changes in operating conditions.

In this way, prediction models are trained only once using historical data from 2014 until end of December 2016 (no measurements are available before this period). The year 2017 (validation set) is used to select the hyper-parameters of the different models. Then, an error metric (performance indicator) is computed between forecasts and real data for each quarter of an hour of January and February 2018 (test set), and the results are represented in Table I. Consequently, the training, validation and test sets are respectively composed of 103872, 35040 and 5472 data points. This results (for the encoder-decoder model) into a training time around 20 minutes (for one configuration). Once the model is trained, its online utilization time (for generating the predictions that are integrated into the robust optimization tool) is around 10^{-6} s.

When dealing with probabilistic forecasting, two (potentially conflicting) notions are important to consider, i.e., calibration (or reliability) and sharpness [59]. Calibration refers to the statistical correctness between the predicted quantiles and the true distribution (i.e., actual observations). Sharpness is a simple measure of the concentration (width) of the predictive distributions. Powerful forecasts must thus find a trade-off between maximizing the sharpness (concentrated intervals), while ensuring that the reliability of the predicted intervals is preserved. To comprehensively assess this compromise between reliability and sharpness, two probabilistic scoring tools are employed.

Firstly, we use the pinball loss E_N, i.e., the mean of the pinball losses, averaged across all quantiles q of interest (in this paper, $q = 5, 15, 25, 35, 45, 50, 55, 65, 75, 85$ and 95%), and throughout the length of the test set ($n = 5472$).

$$E_N = \frac{1}{n}\sum_{t=1}^{n}\sum_{q\in Q} q\max\left(0, y_t - \hat{y}_t^{(q)}\right) + (1-q)\max\left(0, \hat{y}_t^{(q)} - y_t\right) \tag{11}$$

where $\hat{y}^{(q)}$ are the forecasted quantiles, and y the actual observations of the system imbalance (in MW). A lower E_N score indicates a better probabilistic forecast.

However, by averaging all quantiles in the final score, the pinball loss may hide low reliability levels for extreme quantiles. For instance, even if the 5% quantile forecasts can completely fail, it may have a very limited impact on the total score. To address this issue, the Winkler score is used since it is intrinsically associated with a confidence interval [60]. In this way, for a confidence interval of $(1-\alpha)$. 100%, the Winkler score is

[1] A random forest is an ensemble model in which different random trees (i.e., learners) are constructed independently using a different subsample of the data. For prediction tasks, the output is the averaged sum of all trees.

TABLE II
COMPARISON OF DIFFERENT PREDICTION TOOLS FOR JANUARY AND FEBRUARY 2018 ON DIFFERENT METRICS

Topology	Pinball loss	Winkler score					Empirical coverage				
		α=0.1	α=0.3	α=0.5	α=0.7	α=0.9	q=0.05	q=0.25	q=0.5	q=0.75	q=0.95
Probabilistic persistence method	282 MW	426	313	253	209	173	4.3	24.3	51	75.8	94.6
Probabilistic ARIMA	245 MW	360	269	219	182	152	6.2	27	52.5	75	94.2
Quantile regression forest (QRF)	223 MW	332	246	200	166	137	5.4	24.8	51.9	77.6	95.5
MLP	207 MW	310	227	185	154	127	5.6	28	51.2	75.7	95.6
3-MLP	199 MW	299	220	178	148	123	6.8	26.4	48.2	71.4	93.7
GRU-MLP	196 MW	291	216	175	145	120	4.6	24	47.5	75.5	96.1
XGBoost	189 MW	309	212	168	138	114	6.5	28.4	51.4	75.1	94.6
Encoder-decoder	179 MW	268	197	159	132	109	5.4	25.7	52.7	74.4	94.1

defined as:

$$\text{Winkler} = \frac{1}{n}\sum_{t=1}^{n}\begin{cases}\delta_t & L_t \le y_t \le U_t \\ \delta_t + 2\left(L_t - y_t\right)/\alpha & y_t < L_t \\ \delta_t + 2\left(L_t - U_t\right)/\alpha & y_t < U_t\end{cases} \tag{12}$$

where L_t and U_t are respectively the lower and upper bounds of the prediction interval (defined by the confidence level α), $\delta_t = U_t - L_t$ is the predicted interval width. In this way, if the actual observation falls into the predicted interval, the Winkler score is a direct measure of sharpness (and a lower score indicates a better probabilistic forecast). Otherwise, a penalty term is added if an actual value lies outside the constructed interval. The penalty value depends on the severity of the forecast error, and is thus a measure of calibration.

In this paper, the Winkler score is calculated for confidence levels α that correspond to the different risk-attitudes that will be tested into the subsequent robust optimization procedure, i.e., $\alpha = 0.1, 0.3, 0.5, 0.7$ and 0.9. This allows confronting the prediction accuracy in regards with the final purpose of the proposed forecast-driven optimization methodology.Besides, these two metrics are supported by an empirical measure of the coverage of prediction intervals. To that end, we compute (over the test set) the actual percentage of observations below the quantile q, and we compare this value to the corresponding quantile nominal probability q.

In Table II, it can be observed that all neural network models strongly outperform both the probabilistic persistence method and the ARIMA statistical model. It shows that such simple models are poorly suited for modeling the nonlinear behavior (with quick variations) of the system imbalance. Random forests are providing a more reliable framework (that can be implemented in a short period of time), but they also lack the processing power of more complex models.

Another interesting result is that the 3-MLP (3 hidden layers) obtains better results than the shallow MLP, which highlights the importance of depth for leveraging the full potential of neural networks. Moreover, in accordance with [47], the hybrid GRU-MLP model has comparable performances than the simpler 3-MLP model, which illustrates the difficulty to design tailored architectures that improve the prediction accuracy. However, the proposed encoder-decoder, whose temporal nature is designed to optimally exploit both dynamic and static variables, seems to achieve this objective by outperforming all other neural networks for the challenging task of predicting the one-step ahead system imbalance.

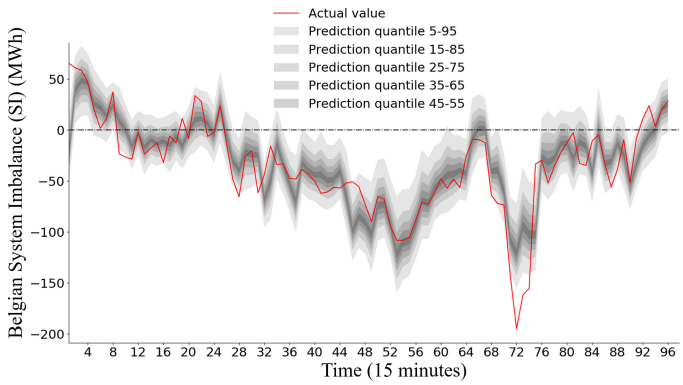

Fig. 5. Probabilistic forecasts of the Belgian system imbalance performed sequentially for the 96 quarters-of-an-hour of the 8th January 2018.

Interestingly, XGBoost yields similar results than the MLP (according to the Winkler score) for extreme quantiles, which may be harmful when predictions are intended to be embedded within risk-aware optimization strategies. However, its accuracy increases for large confidence levels α (i.e., $\alpha = 0.3, 0.5, 0.7$ and 0.9), which emphasizes its good ability to predict values close to the mean.

It should also be noted that the empirical coverage (computed over a representative test set of 5472 points) shows that all tools give consistent quantiles (e.g., 94.1% of the actual system imbalances are below the predicted quantile $q = 0.95$ for the encoder-decoder), which demonstrates the reliability of using quantile regression (7) for probabilistic forecasting.

For illustrating the quality of the results obtained using the encoder-decoder network with the (non-parametric) quantile loss function, probabilistic forecasts of the system imbalance carried out sequentially (one step-ahead) for the 96 quarters-of-an-hour of the 8th January 2018 are shown in Fig. 5.

One can see that the predicted intervals tend to properly encompass the actual realizations of uncertainties (i.e., the volatility of the system imbalance is well captured). More interestingly, we observe for certain time periods (e.g., around quarters of an hour 4 and 72) that the prediction intervals are sufficiently sharp so that all quantiles are positive or negative. Those moments correspond thus to ideal situations for deviating from the balanced position (with the necessary security margins)

TABLE III
IMPACT OF DIFFERENT FORECASTERS AND DIFFERENT RISK-ATTITUDE STRATEGIES IN THE PARTICIPATION IN THE SINGLE PRICE IMBALANCE SETTLEMENT

	Aggregated expected profits (ex-ante) over the test set [k€]						Aggregated actual profits (ex-post) over the test set [k€]					
Perfect forecast	4201						4201					
	Bounds 05-95	Bounds 15-85	Bounds 25-75	Bounds 35-65	Bounds 45-55	Bound 50	Bounds 05-95	Bounds 15-85	Bounds 25-75	Bounds 35-65	Bounds 45-55	Bound 50
QRF	161	554	1047	1624	2316	2722	480	939	1131	1019	634	245
MLP	359	882	1396	1859	2533	2975	850	1126	1203	1015	637	264
3-MLP	587	1222	1704	2258	2784	3136	1040	1282	1184	971	564	332
GRU-MLP	362	949	1539	2213	2773	3130	864	1267	1208	1075	772	411
XGBoost	216	954	1574	2111	2697	3070	798	1247	1389	1228	950	663
Encoder-decoder	574	1213	1765	2315	2959	3315	1151	1429	1391	1251	1023	847

to make profit in the imbalance settlement. This observation will be verified and quantified in the next Subsection IV.B, where the probabilistic forecasts will be fed into a robust optimization tool.

B. Numerical Results of the Real-Time Market Optimization

In this work, an aggregated battery storage system is studied, whose technical parameters are borrowed from [61]. This energy storage system (ESS) is characterized by a symmetric and continuous variation range for the output power (120 MW), with good ramping (60 MW/min) and energy capabilities (240 MWh). The power limitations (8e)–(8f) and technology-specific constraints (8g) associated with the ESS can be found in the Appendix (Section VI). The costs C^+ and C^- are 50 and 30 € /MWh respectively, and ensure that the unit will exploit only profitable price spreads between the charged and discharged energy, accounting for the operating costs and efficiency losses of the unit [62]. Practically, the ESS will adopt a long position only if $\lambda^{SI}>50$ €/MWh, and a short position only if $\lambda^{SI}<30$ €/MWh.

We consider that no contracts were agreed in advance, and that the whole flexibility is thus available in real-time. The resulting MILP model is implemented in Python, using the Pulp library, and solved with CPLEX, on an Intel Core i7-2630 CPU @ 2.0 GHz with 8GB RAM.

The developed procedure (probabilistic forecasts of the system imbalance, which are then used into the robust optimization tool) is run sequentially over the entire test set horizon (at the start of each quarter-of-an-hour of January and February 2018). The storage system is initially half-charged, but this value is adjusted between consecutive time steps based on the real-time dispatch of the unit.

An ex-post out-of-sample analysis is then carried out, which consists in confronting, at each quarter-of-an-hour, the decisions $e^{imb,+}$ and $e^{imb,-}$ (obtained at the end of the optimization) with the actual realizations of the system deviation. This allows computing the actual (ex-post) profit of the market player, which allows to quantify the value of advanced predictions on actual field operations. This analysis is performed in parallel for different configurations, in order to quantify the impact on the operational profit of: (i) improved forecasts of the system imbalance, and (ii) the risk-attitude strategy (which is defined through different bounds of the uncertainty set, taken from the different quantiles of the probabilistic forecasts). Practically, the outcomes from each prediction tool for each risk-policy are incorporated into the robust dispatch strategy, and the quality of the resulting decisions is compared in Table III.

As a measure of the upper bound of the profit that can be generated, the optimization problem was firstly run with the perfect knowledge of the system imbalance. Then, the aggregated profits (summed over all quarters-of-an-hour of the 2-months test set period) that were expected at the end of the optimization procedure (left part of Table III) are put into perspective with the actual profits given by the (ex-post) clearing of the imbalance settlement, based on the optimized position $e^{imb,+}$ and $e^{imb,-}$ of the market participant. From these results, several trends can be identified.

Firstly, it is important to mention that the computational time is always lower than 1 minute (whereas the forecasting tool necessitates less than 1 second to provide the probabilistic forecasts of interest), which enables our robust formulation with probabilistic guarantees to be exploited in the real-time operation of market players with flexible resources.

Secondly, it is observed that the quality of predictions constitutes a prominent factor to take reliable decisions. In this way, using the risk-optimal predictions from our encoder-decoder model is here associated with an increase of profit between 40 k€ (compared to the second best model, i.e., XGBoost) and 298 k€ (for quantile regression forest), which corresponds to relative increases of respectively 2.8% and 21% throughout the test period of January-February 2018. In this way, even small improvements in the prediction accuracy can result into significant additional profits. This strongly paves the way to further research to enhance prediction tools. In this way, the best predictor yields a risk-optimal profit of 1429 k€ , whereas perfect forecasts would generate 4201 k€ . Moreover, in accordance with its forecast (Section IV.A), XGBoost is very competitive for risky strategies, but leads to lesser performance for conservative approaches (since its extreme quantiles are then less accurate).

Thirdly, the robust formulation ensures that the portfolio never participates in the imbalance settlement when the bounds of the uncertainty set (i.e., the selected quantiles of the predicted system imbalance) are of different signs. On the contrary, when the predicted bounds are sign-consistent, the market player may deviate from its balanced position with the objective to help at restoring the power system balance. For this to happen, the flexibility margins have to be still available in real-time, and the imbalance price must be sufficiently attractive to cover all intrinsic costs of the unit (which is true in most cases due to the extreme nature of the imbalance settlement). In this way, in the

TABLE IV
COMPARISON OF THE PROFIT GENERATED BY PRICE-MAKER AND PRICE-TAKER ASSUMPTIONS FOR DIFFERENT RISK-ATTITUDES

	Aggregated expected profits over the test set [k€]		Aggregated ex-post profits over the test set [k€]	
	Price-taker	Price-maker	Price-taker	Price-maker
Bounds 05-95	1468	574	773	1151
Bounds 15-85	3142	1213	-394	1429
Bounds 25-75	4370	1765	-1588	1391
Bounds 35-65	5458	2315	-2870	1251
Bounds 45-55	6598	2959	-4396	1023
Bound 50	7296	3315	-5147	847

risk-neutral strategy (where a single bound of the uncertainty set is considered), the market player almost systematically participates in the imbalance settlement, which represents an amount of 75255 MWh over the test set period (when the encoder-decoder is used as forecaster). By contrast, the associated risk-averse policy (05–95 quantiles) leads to a moderate contribution of 16352 MWh (which are played 27% of the time, i.e., an average of 25 times a day for the 96 daily settlement periods).

Fourthly, we see that the expected and actual profits can be very different. Such discrepancies are exacerbated for risky attitudes since the decisions are then based on forecasts with low reliability. Such risky approaches tend to overestimate the profit that will be actually generated, leading to ex-post disappointment when the actual outcome is revealed. For instance, the approach with the 45–55 quantiles represents a reliability of 10% that the actual system imbalance lies in the prediction interval, which ultimately jeopardizes the performance of the optimization. This illustrates the need to rely on risk-aware approaches to hedge against such situations.

In our case, the risk-optimal approach is the optimization performed with the 15–85 quantiles (obtained with the encoder-decoder), since this tool leads to the highest actual profits, and therefore to the best tradeoff between conservativeness and economic performance. This strategy is sufficiently audacious to properly take advantage of favorable situations, while hedging against the market volatility, by avoiding to participate when the conditions are unsure. Indeed, results highlight that this risk-optimal policy leads to erroneous offers (that infer financial penalties) occurring 6.5% of the time. In comparison, the deterministic (quantile 50) and risk-seeking policies (45–55, 35–65 and 25–75 quantiles) result in respectively 38%, 32%, 22% and 14% of erroneous offers, while such misinformed decisions happen respectively 1% of the time for the more risk-averse strategies (05–95 quantiles).

Finally, we study the impact of the price-maker assumption on the performance of the dispatch procedure. To that end, the optimization is performed in the same conditions (based on the predictions given by the encoder-decoder), in a price-taker setting. Practically, the lower-level (9) is cleared independently (disregarding the actions of the market player), and the resulting imbalance price is then treated as an exogenous variable in the subsequent (single-stage) optimization. The outcomes (in terms of expected and actual profits) for different risk-attitudes are shown in Table IV.

Results indicate that the price-taker assumption is not appropriate to model the very-short-term participation of market players in the imbalance settlement. Indeed, when the forecasted quantiles of the system imbalance are of the same sign, the actor provides its full capacity (regardless of its impact on the system state), which often results in switching the system conditions. The ESS then charges at high prices, and discharges for low prices, which ultimately leads to negative profit (especially for risky policies). In this way, the price-taker assumption is systematically over-optimistic, which results in very inefficient strategies when the actual outcome (profit) is revealed. The price-maker assumption, on the other hand, allows to better hedge against the real-time volatility of the power system conditions, by properly considering the actions of the market player.

V. CONCLUSION

In this paper, we propose an integrated forecast-driven methodology for the risk-aware participation in the single price imbalance settlement of European electricity markets. Exploiting this real-time market opportunity is challenging due to the extreme variability and very limited predictability of the system conditions. In order to incorporate this system uncertainty into the decision procedure, we have developed improved probabilistic predictions of the system imbalance, based on a promising deep-learning model named encoder-decoder. It should be noted that such predictions of the system imbalance could also be useful for the system operator in order to help him in activating the adequate amount of operating reserves. Results demonstrate that the temporal nature of the encoder-decoder tool allows to generate more accurate (tightened) quantiles in comparison with the other state-of-the-art models, which ultimately results in a more efficient (informed) participation in the imbalance settlement.

In a second step, these probabilistic forecasts are used as inputs into the subsequent decision procedure, allowing to optimize the risk-attitude through different characterizations of the uncertainty set. In parallel, fast actions are required to manage such an unstable environment. A computationally efficient decision-making process is therefore implemented, based on robust optimization, in order to efficiently capture market opportunities and maximize the player's profit in a price-maker setting (to account of his market power). Interestingly, it has been shown that the profit can be significantly increased, not only by relying on an optimal policy to manage and optimize the risk, but also by improving (even slightly) the quality of the underlying predictions.

As a perspective, this near real-time formulation could be integrated into the day-ahead horizon framework such that the remaining flexibility margins available in real-time are fully optimized with regard to other previous market opportunities.

APPENDIX

This appendix contains the technology specific constraints associated with the aggregated battery storage system. This energy storage system (ESS) can deviate from its balanced position in four different ways [63]:

- Upward deviation (i.e., positive imbalance), either by reducing the charging power $\Delta p_{\upsilon}^{c,+}$, or by increasing the discharging power $\Delta p_{\upsilon}^{d,+}$.
- Downward deviation (i.e., negative imbalance), either by increasing the charging power $\Delta p_{\upsilon}^{c,-}$, or by increasing the discharging power: $\Delta p_{\upsilon}^{d,-}$.

These ESS-based power deviations are constrained by the power limitations are constrained by the power limitations and the initially planned (scheduled) output of the ESS for the considered quarter-of-an-hour:

$$0 \leq \Delta p_{\upsilon}^{c,+} \leq P_{\upsilon}^{c}, \quad \forall \upsilon \tag{A1}$$

$$0 \leq \Delta p_{\upsilon}^{c,-} \leq P_{\upsilon}^{\max} - P_{\upsilon}^{c}, \quad \forall \upsilon \tag{A2}$$

$$0 \leq \Delta p_{\upsilon}^{d,-} \leq P_{\upsilon}^{d}, \quad \forall \upsilon \tag{A3}$$

$$0 \leq \Delta p_{\upsilon}^{d,+} \leq P_{\upsilon}^{\max} - P_{\upsilon}^{d}, \quad \forall \upsilon \tag{A4}$$

where P_{υ}^{c} and P_{υ}^{d} are the scheduled power of ESS υ in charging and discharging operating modes respectively, while $P_{\upsilon}^{\max}$ is the installed capacity (in MW).

Then, limits on the energy content (A5) and (A6) have to be respected at each time step. The state-of-charge (*SOC*) is updated at each time step based on previous actions, and accounts for energy losses originating from charge and discharge inefficiencies (ε_{υ} is the round trip efficiency of the ESS). The energy content limitations ($E_{\upsilon,t}^{\min}, E_{\upsilon,t}^{\max}$) can be adapted for each quarter-of-an hour *t* to guarantee that sufficient energy can be stored or discharged by the ESS to meet the needs associated with the following time steps.

$$SOC_{\upsilon,t} - \Delta t \left(\Delta p_{\upsilon}^{c,+} \sqrt{\varepsilon_{\upsilon}} + \frac{\Delta p_{\upsilon}^{d,+}}{\sqrt{\varepsilon_{\upsilon}}} \right) \geq E_{\upsilon,t}^{\min}, \quad \forall \upsilon, t \tag{A5}$$

$$SOC_{\upsilon,t} + \Delta t \left(\Delta p_{\upsilon}^{c,-} \sqrt{\varepsilon_{\upsilon}} + \frac{\Delta p_{\upsilon}^{d,-}}{\sqrt{\varepsilon_{\upsilon}}} \right) \leq E_{\upsilon,t}^{\max}, \quad \forall \upsilon, t \tag{A6}$$

where Δt is the time period (here, 0.25 hour). Finally, constraint (A7) ensures that real-time power deviations satisfy the ESS's ramping limits:

$$\Delta p_{\upsilon}^{c,+}, \Delta p_{\upsilon}^{d,+}, \Delta p_{\upsilon}^{c,-}, \Delta p_{\upsilon}^{d,-} \leq 60 \Delta t \cdot R_{\upsilon}, \quad \forall \upsilon \tag{A7}$$

where $R_{\upsilon,t}$ is the available ramping ability (in MW/min).

ACKNOWLEDGMENT

The authors would like to thank the Energy Transition Funds Project (EPOC 2030-2050) organized by the FPS Economy, S.M.E.s, Self-Employed and Energy for providing financial support.

REFERENCES

[1] IEA, "Status of power system transformation 2018: Advanced power plant flexibility," Paris, France, 2018.

[2] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, and L. F. Ochoa, "Evaluating and planning flexibility in sustainable power systems," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 200–209, Jan. 2013.

[3] J. Garcia-Gonzalez, R. de la Muela, L. Santos, and A. Gonzalez, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 460–468, May 2008.

[4] J. Matevosyan and L. Söder, "Minimization of imbalance cost trading wind power on the short-term power market," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1396–1404, Aug. 2006.

[5] J. Bottieau, F. Vallée, Z. De Grève, and J.-F. Toubeau, "Leveraging provision of frequency regulation services from wind generation by improving day-ahead predictions using LSTM neural networks," in *Proc. IEEE Int. Energy Conf.*, Limassol, 2018, pp. 1–6.

[6] A. J. Conejo, M. Carrión, and J. M. Morales, *Decision Making Under Uncertainty in Electricity Markets*. New York, NY, USA: Springer, 2012.

[7] M. Kazemi, H. Zareipour, N. Amjady, W. D. Rosehart, and M. Ehsan, "Operation scheduling of battery storage systems in joint energy and ancillary services markets," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1726–1735, Oct. 2017.

[8] J.-F. Toubeau, Z. De Grève, and F. Vallée, "Medium-term multimarket optimization for virtual power plants: A stochastic-based decision environment," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1399–1410, Mar. 2018.

[9] F. Bourry, L. Costa, and G. Kariniotakis, "Risk-based strategies for wind/pumped-hydro coordination under electricity markets," in *Proc. IEEE Power Tech Conf.*, Bucharest, Romania, Jun. 28–Jul. 2, 2009, pp. 1–8.

[10] S. I. Vagropoulos and A. G. Bakirtzis, "Optimal bidding strategy for electric vehicle aggregators in electricity markets," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4031–4041, Nov. 2013.

[11] H. Ding, Z. Hu, and Y. Song, "Rolling optimization of wind farm and energy storage system in electricity markets," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2676–2684, Sep. 2015.

[12] H. Ding, P. Pinson, Z. Hu, and Y. Song, "Integrated bidding and operating strategies for wind-storage systems," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 163–172, Jan. 2016.

[13] A. Hellmers, M. Zugno, A. Skajaa, and J. M. Morales, "Operational strategies for a portfolio of wind farms and CHP plants in a two-price balancing market," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2182–2191, May 2016.

[14] R. A. C. van der Veen, A. Abbasy, and R. A. Hakvoort, "Agent-based analysis of the impact of the imbalance pricing mechanism on market behavior in electricity balancing markets," *Energy Econ.*, vol. 34, pp. 874–881, 2012.

[15] ACER, "Recommendation of the agency for the cooperation of energy regulators No 03/2015–on the network code on electricity balancing," New Taipei City, Taiwan, 2015.

[16] CREG, "Study on the functioning and price evolution of the Belgian wholesale electricity market–monitoring report 2017," Xi'an, China, 2018.

[17] M. Zugno, J. Morales, P. Pinson, and H. Madsen, "Pool strategy of a price-maker wind power producer," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3440–3450, Aug. 2013.

[18] L. Baringo and A. J. Conejo, "Strategic offering for a wind power producer," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4645–4654, Nov. 2013.

[19] A. Abdisalaam, I. Lampropoulos, J. Frunt, G. Verbong, and W. L. Kling, "Assessing the economic benefits of flexible residential load participation in the Dutch day-ahead spot and balancing markets," in *Proc 9th Conf. Eur. Energy Market*, Florence, Italy, May 10–12, 2012, pp. 1–8.

[20] J. Zapata Riveros, R. Donceel, J. Van Engeland, W. D'Haeseleer, "A new approach for near real-time micro-CHP management in the context of power system imbalances—A case study," *Energy Convers. Manage.*, vol. 89, pp. 270–280, 2015.

[21] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *Int. J. Forecast.*, vol. 30, no. 2, pp. 357–363, Apr.–Jun. 2014.

[22] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable Sustain. Energy Rev.*, vol. 32, no. 5, pp. 255–270, 2014.

[23] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renewable Sustain. Energy Rev.*, vol. 81, pp. 1548–1568, 2018.

[24] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecast.*, vol. 32, pp. 914–938, 2016.

[25] J. Juban, L. Fugon, and G. Kariniotakis, "Probabilistic short-term wind power forecasting based on kernel density estimators," in *Proc. Eur. Wind Energy Conf.*, Milan, Italy, 2007, pp. 1–11.

[26] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problem," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496–1503, Aug. 2010.

[27] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.

[28] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation—With application to solar energy," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850–3863, Sept. 2016.
[29] C. Wan, J. Lin, Y. Song, Z. Xu, and G. Yang, "Probabilistic forecasting of photovoltaic generation: An efficient statistical approach," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2471–2472, May 2017.
[30] Z. Shi, H. Liang, and V. Dinavahi, "Direct interval forecast of uncertain wind power based on recurrent neural networks," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1177–1187, Jul. 2018.
[31] J. Dowell and P. Pinson, "Very-short-term probabilistic wind power forecasts by sparse vector autoregression," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 763–770, Mar. 2016.
[32] A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, "A new fuzzy-based combined prediction interval for wind power forecasting," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 18–26, Jan. 2016.
[33] R. Bessa, V. Miranda, A. Botterud, J. Wang, and E. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660–669, Oct. 2012.
[34] Y. Zhang and J. Wang, "K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting," *Int. J. Forecasting*, vol. 32, no. 3, pp. 1074–1080, Jul. 2016.
[35] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2767–2778, Jul. 2017.
[36] K. Hatalis, A. J. Lamadrid, K. Scheinberg, and S. Kishore, "Smooth pinball neural network for probabilistic forecasting of wind power," 2017, *arXiv:1710.01720*.
[37] K. Bruninx and E. Delarue, "Endogenous probabilistic reserve sizing and allocation in unit commitment models: Cost-effective, reliable, and fast," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2593–2603, Jul. 2017.
[38] E. Nasrolahpour, J. Kazempour, H. Zareipour, and W. D. Rosehart, "A bilevel model for participation of a storage system in energy and reserve markets," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 582–598, Apr. 2018.
[39] J.-F. Toubeau, J. Bottieau, F. Vallee, and Z. De Greve, "Deep learning-based multivariate probabilistic forecasting for short-term scheduling in power markets," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1203–1215, Mar. 2019.
[40] D. Bertsimas, D. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Rev.*, vol. 53, pp. 464–501, 2011.
[41] M. P. Garcia and D. S. Kirschen, "Forecasting system imbalance volumes in competitive electricity markets," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 240–248, Feb. 2006.
[42] M. Khodayar and J. Wang, "Spatio-temporal graph deep neural network for short-term wind speed forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 670–681, Apr. 2019, doi: 10.1109/TSTE.2018.2844102.
[43] J.-F. Toubeau, J. Bottieau, F. Vallée, and Z. De Grève, "Improved day-ahead predictions of load and renewable generation by optimally exploiting multi-scale dependencies," in *Proc. 7th IEEE Conf. Innovative Smart Grid Technol.*, Dec. 2017, pp. 1–5.
[44] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[45] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Proc.*, 2014, pp. 1724–1734.
[46] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, *arXiv:1412.3555*.
[47] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Appl. Energy*, vol. 221, pp. 386–405, 2018.
[48] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-Decoder approaches," 2014, *arXiv:1409.1259*.
[49] Y. Wang, D. Gan, M. Sun, N. Zhang, Z. Lu, and C. Kang, "Probabilistic individual load forecasting using pinball loss guided LSTM," *Appl. Energy*, vol. 235, pp. 10–20, 2019.
[50] P. J. Huber and E. Ronchetti, *Robust Statistics*, 2nd ed. Hoboken, NJ, USA: Wiley, 2013.
[51] Elia Group, 2018. [Online]. Available: http://www.elia.be/en/grid-data/balancing/available-regulation-capacity.
[52] J. Morales, A. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewable in Electricity Markets*. New York, NY, USA: Springer, 2014.
[53] Elia Group, "Elia grid data," 2019. [Online]. Available: www.elia.be/en/grid-data/data-download
[54] J. Bottieau, L. Hubert, Z. De Grève, F. Vallée, and J.-F. Toubeau, "Historical database for very-short-term probabilistic forecasting of system imbalance," 2019. [Online]. Available: https://doi.org/10.5281/zenodo.2668485
[55] N. Meinshausen, "Quantile regression forests," *J. Mach. Learn. Res.*, vol. 7, pp. 983–999, 2006.
[56] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.
[57] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Proc. 24th Int. Conf. Neural Inf. Process. Syst.*, 2011, pp. 2546–2554.
[58] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–41.
[59] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. Roy. Statistical Soc. B*, vol. 69, pp. 243–268, 2007.
[60] R. L. Winkler, "A decision-theoretic approach to interval estimation," *J. Amer. Statist. Assoc.*, vol. 67, no. 337, pp. 187–191, 1972.
[61] E.G. Kardakos, C.K. Simoglou, and A.G. Bakirtzis, "Optimal offering strategy of a virtual power plant: A stochastic bi-level approach," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 794–806, Mar. 2016.
[62] A. Belderbos, E. Delarue, and W. D´haeseleer, "Calculating the levelized cost of electricity storage," in *Proc. IAEE Int. Conf. Energy*, Bergen, Norway, 2017.
[63] J.-F. Toubeau, Z. De Grève, P. Goderniaux, F. Vallée, and K. Bruninx, "Chance-constrained scheduling of underground pumped hydro energy storage in presence of model uncertainties," *IEEE Trans. Sust. Energy*, to be published, doi: 10.1109/TSTE.2019.2929687.

Jérémie Bottieau (S'17) received the Diploma in electrical engineering from the University of Mons, Mons, Belgium, where he has been working toward the Ph.D. degree with the Department of Electrical Power Engineering since 2017. His research interests include short-term forecasting and optimization in electricity markets.

Louis Hubert is currently working toward the master's degree at the Department of Electrical Power Engineering, the University of Mons, Mons, Belgium.

Zacharie De Grève (M'12) received Electrical and Electronics Engineering degree in 2007 from the Faculty of Engineering of Mons, University of Mons, Mons, Belgium, where he received the Ph.D. degree in electrical engineering in 2012. He was a Research Fellow with the Belgian Fund for Research (F.R.S/FNRS) until 2012. He is currently a Researcher with the Electrical Power Engineering Unit, the University of Mons, and a part-time Lecturer since September 2019. He conducts transverse research in machine learning, optimization and energy economics, applied to modern electricity networks with a high share of renewables, in order to contribute to the energy transition.

François Vallée (M'09) received the degree in civil electrical engineering and the Ph.D. degree in electrical engineering from the Faculty of Engineering, University of Mons, Mons, Belgium, in 2003 and 2009, respectively. He is currently an Associate Professor with the Electrical Power Engineering Unit, Faculty of Engineering of Mons, University of Mons. He has authored or coauthored of several publications and his Ph.D. work has been awarded by the SRBE/KBVE Robert Sinave Award in 2010. His research interests include PV and wind generation modeling for electrical system reliability studies in presence of dispersed generation. He is currently a member of the Governing Board from the 'Société Royale Belge des Electriciens - SRBE/KBVE' (2017) and an Associate Editor of the *International Transactions on Electrical Energy Systems* (Wiley).

Jean-François Toubeau (M'18) was born in Mons, Belgium, in 1990. He received the degree in civil electrical engineering and the Ph.D. degree in electrical engineering from the University of Mons, Mons, Belgium, in 2013 and 2018, respectively. He is currently a Postdoctoral Researcher with the Belgian Fund for Research (F.R.S/FNRS) within the Electrical Power Engineering Unit, Faculty of Engineering of Mons, University of Mons. His research interests include decision making in the context of power markets as well as data analytics.