Probabilistic State Forecasting and Optimal Voltage Control in Distribution Grids under Uncertainty

Thierry Zufferey, Sandro Renggli, and Gabriela Hug
Power Systems Laboratory, ETH Zurich, Switzerland
{thierryz, srenggli, hug}@ethz.ch

***Abstract*—Nowadays, the uncertainty in distribution systems rises, notably due to an increasing share of solar panels and electric vehicles whose power production and consumption are characterized by a high volatility. This poses challenges to distribution system operators to ensure stable and secure operation of their grid. Hence, an optimal integration of these distributed energy resources in real-time control schemes inevitably relies on appropriate forecasts of the near-future system state. This paper investigates the short-term probabilistic state prediction of low-voltage grids for operation purposes. The performance of two quantile forecasting algorithms is evaluated for different levels of distributed energy resources penetration and availability of measurements. Quantile forecasts are finally integrated into the framework of an optimization problem that aims at minimizing the costs associated with overvoltages by suitable solar power curtailment. The advantages of quantile forecasts considering different imbalance prices are demonstrated.**

***Index Terms*—distributed energy resources, low-voltage distribution grid, optimal power flow, quantile forecasting, voltage control**

I. Introduction

The share of Distributed Energy Resources (DERs) in today's distribution grids is increasing. These DERs, mainly in the form of Photovoltaic (PV) systems and Electric Vehicles (EVs), introduce new operational challenges to Distribution System Operators (DSOs). On the one hand, PV systems synchronously inject active power into the grid according to the volatile solar irradiance and can considerably increase the voltage in the system. On the other hand, EVs represent stochastic loads which consume substantial amounts of active power and decrease the voltage during the charging phase. Since DSOs are responsible for the safe operation of their networks by avoiding voltage band violations or line overloadings, it becomes crucial to properly observe and foresee voltages and power flows down to the low-voltage level.

Nowadays, the wide roll-out of Smart Meters (SMs) allows DSOs to gain insights into their system, notably by using Distribution System State Estimation (DSSE) techniques. Widely discussed in the current literature [1], Weighted Least Square (WLS) based algorithms perform particularly well on the transmission grid level but cannot properly exploit highly volatile measurements in distribution grids [2]. Alternatively, Machine Learning (ML) algorithms can be used to estimate the most likely state of a system. The authors in [3] propose a closed-loop approach where a WLS algorithm is fed by the load forecasts of a Non-linear Auto-Regressive Exogenous (NARX) model. K-Nearest Neighbor (KNN) based approaches are presented in [2], [4], where the actual state is estimated by using historical system states whose corresponding measurements are similar to the real-time measurements. Nevertheless, state-of-the-art DSSE approaches are deterministic and cannot give insight into the state estimation errors which can largely vary between time steps.

While the errors do not follow a given distribution, the quantification of the near-future state uncertainty is essential in distribution grids populated with highly volatile loads and DERs. Probabilistic forecasting algorithms enable a comprehensive prediction by covering the entire uncertainty range without assumption on the error distribution and by updating their prediction in real time [5]. For example, the authors in [6], [7] elaborate on regression Neural Networks (NNs) and Long Short-Term Memory (LSTM) algorithms for quantile load forecasting with promising results. However, the literature mainly focuses on load forecasting and disregards the direct prediction of other quantities such as voltage and line loading which are relevant for DSOs. In this paper, a quantile NN and a novel probabilistic version of the KNN algorithm are designed to predict net power consumptions, line power flows, and bus voltage magnitudes. Their performance is evaluated on a real low-voltage distribution system of the City of Basel in Switzerland, considering multiple levels of PV and EV penetration. In addition, we study the added value of real-time instead of time-delayed SM measurements, and of an additional feature that indicates at which points in time the EVs are charging.

Moreover, in the literature on power systems optimization, a focus has been recently given to control schemes that explicitly consider uncertainty to cope with the volatile and hardly predictable nature of DERs [8]. This is notably the case for chance-constrained Optimal Power Flow (OPF) [9]. However, they usually assume a given error distribution and overlook the large temporal variations in uncertainty, which does not reflect the very volatile reality. Therefore, this paper promotes the use of short-term quantile forecasts by directly integrating them into a two-stage OPF control scheme. It aims at optimally estimating in advance the required PV power curtailment to keep the voltages within limits and avoid expensive curtailment in real time.

The remainder of this paper is structured as follows. Section II presents the quantile forecasting algorithms and Sec. III details multiple OPF schemes for voltage control. The grid

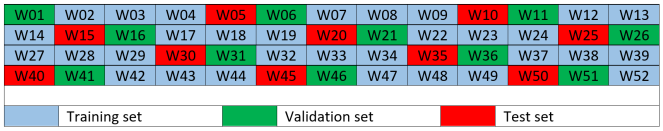

Fig. 1. Weeks of the data set split into training, validation and test set

model used in the case study and the associated measurements are described in Sec. IV whereas the performance of the forecasting algorithms and of the control schemes is discussed in Sec. V. Section VI summarizes the main outcomes and outlines future work.

II. Probabilistic State Forecasting

While still widely used in the literature on control of active distribution grids, perfect forecasts are unrealistic while deterministic forecasts cannot account for the underlying uncertainty [10]. Moreover, the assumption of a given error distribution gives a biased image of the reality [8]. In this paper, we promote very short-term quantile forecasting algorithms which are probabilistic methods that quantify the prediction uncertainty for each time step.

A. Setup of quantile forecasting

In this paper, active and reactive power consumption P_{cons} and Q_{cons}, active and reactive power flow P_{flow} and Q_{flow}, respectively, and bus voltage magnitude V are predicted one hour ahead by a quantile NN and a novel probabilistic version of KNN. The different quantities and quantiles are correlated and could be predicted simultaneously. However, a preliminary study suggests that the use of one single ML model per grid component, per quantity and per quantile is more efficient in terms of computation time and accuracy, as long as the most influencing quantities belong to the feature set. For each algorithm, the corresponding hyperparameters are selected by grid search. For each ML model, the data set over one full year is split into training, validation and test sets according to Fig. 1. This partitioning accounts for the seasonal behavior of measurements while yielding continuous test sets used in the control schemes described in Sec. III. The training process consists of minimizing the pinball loss function, which creates separate forecasts for different quantiles:

$$J_q = \frac{1}{N}\sum_{n=1}^{N}\begin{cases}(y_n - \hat{y}_{q,n})\,q & \text{if } y_n \geq \hat{y}_{q,n},\\ (\hat{y}_{q,n} - y_n)\,(1-q) & \text{if } y_n < \hat{y}_{q,n},\end{cases} \quad (1)$$

where q is the target quantile, y_n is the target value of sample n, $\hat{y}_{q,n}$ is the prediction for quantile q of sample n, and N is the number of samples in the training set. The loss function is asymmetric such that for any quantile higher (or lower) than the 50%-quantile, forecasting errors due to underestimation of the target value get penalized more (or less) than the errors due to overestimation. Note that the 50%-quantile corresponds to the usual point forecast.

TABLE I
Basic feature set for the prediction of bus quantities

Category	Features	Time Delay
Online	V, P_{cons} and Q_{cons}	-
Recent	V, P_{cons} and Q_{cons}	1 hour and 2 hours
Historical	V, P_{cons} and Q_{cons}	1 day and 2 days
Weather	temperature and solar irradiance	-
Calendar	hour, weekday and holiday flag	-

TABLE II
Basic feature set for the prediction of line quantities

Category	Features	Time Delay
Online	V_1, V_2, P_{flow} and Q_{flow}	-
Recent	V_1, V_2, P_{flow} and Q_{flow}	1 hour and 2 hours
Historical	V_1, V_2, P_{flow} and Q_{flow}	1 day and 2 days
Weather	temperature and solar irradiance	-
Calendar	hour, weekday and holiday flag	-

B. Feature Set

Tables I and II summarize the basic sets of input features for bus quantities (i.e., active and reactive power consumption, voltage magnitude) and line quantities (i.e., active and reactive power flow), respectively. Note that the forecast of a certain quantity also uses the measurements of other quantities to profit from their physical coupling in the grid (e.g., voltages at connected buses are features for the prediction of power line flows). All measurements are standardized and calendar features are one-hot encoded.

In the case study, we investigate the added value of real-time SM data compared to only time-delayed SM data. In the latter case, only day-ahead SM measurements up to midnight are assumed to be accessible. Hence, online and recent features are not available, but the historical feature set is further enhanced by quarter hourly values around the corresponding values one and two days in the past. This ensures a sufficient number of grid measurements in the feature set. Moreover, we assess the impact of knowing the starting time and duration of EV charging events. In practice, this could be inferred from the GPS location and the state of charge of the vehicle[1]. In this case, a binary EV charging feature is added to the buses where an EV is assigned to and to all lines connected to those buses.

C. Quantile Neural Network

NN models are well-known ML algorithms inspired by the structure of the human brain that can approximate any non-linear function via a weighted linear combination of neurons organized in layers. In this study, depending on the forecasted quantity, the NN contains 4 or 6 hidden layers and the neurons are associated with the ELU or RELU activation function. During the training process, the weights of the neural network are adjusted in order to minimize the pinball loss function defined in (1). In addition, L_2-regularization is considered and up to 10% of the neurons can be dropped to prevent overfitting. Note that Lasso regression and Principal Component Analysis have been tested to reduce the feature set dimension but none of the methods is conclusive in terms of forecast accuracy.

[1]This of course raises privacy concerns but is out of scope of this paper.

D. Quantile K-Nearest Neighbor

In contrast to using linear combinations of the input features to find a prediction for the target values, the KNN algorithm assumes that if a similar grid state defined by the input features has been observed at a certain point in the past, the future target value of a grid quantity in one hour from now is similar to the value one hour after the state found in the past. For the purpose of this paper, the similarity measure is based on the weighted Minkowski distance:

$$d_{t,m} = |(\boldsymbol{w}^{\text{mink}})^T (\boldsymbol{x}_m - \boldsymbol{x}_t)|_p, \ \forall m \in \mathcal{M}_t, \tag{2}$$

where $\boldsymbol{x}_m$ and $\boldsymbol{x}_t$ are the feature vectors of past sample m and of current sample t, respectively, $\boldsymbol{w}^{\text{mink}}$ is a vector of weights, $\mathcal{M}_t$ is the set of past samples at current time t, and $p \in \{1, 2\}$ is a hyperparameter defining the norm. The Minkowski weights are set according to the importance of the corresponding features, which is given by Lasso cross-validation in this study. After calculating all Minkowski distances $d_{t,m}$ between the past system states $\boldsymbol{x}_m$, $\forall m \in \mathcal{M}_t$, and the current system state $\boldsymbol{x}_t$, the corresponding target values y_i^{knn} , $\forall i \in \{1, \dots, k\}$, of the k states showing the smallest distances are linearly combined to predict the future target y_t. In order to obtain a suitable number of neighbors k and the optimal KNN weights $\boldsymbol{w}^{\text{knn}}$, ten small optimization problems for $k \in \{5, 10, \dots, 50\}$ are solved:

$$J_{\text{knn}} = \min_{\boldsymbol{w}^{\text{knn}}} \frac{1}{N} \sum_{n=1}^{N} |(\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n|, \tag{3a}$$

$$\text{subject to} \quad w_i^{\text{knn}} \geq 0, \ \forall i \in \{0, \dots, k\}, \tag{3b}$$

where y_n is the target value of sample n, $\boldsymbol{y}_n^{\text{knn}}$ is a vector of targets of the k nearest neighbors of y_n (including a bias term of 1), $\boldsymbol{w}^{\text{knn}}$ is a vector of neighbor weights, and N is the number of samples in the training set. The cost function (3a) is based on the Mean Absolute Error (MAE) function and allows to perform point forecasts. Constraint (3b) guarantees that all neighbors and the bias term have to contribute positively to the result of the linear combination. The training of this algorithm is a proposed extension of the standard KNN algorithm such that quantile forecasts can be obtained by replacing the MAE function by the pinball loss function. More precisely, $|(\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n|$ is replaced by $((\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n)\, q$ if $y_n \geq (\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}}$ and by $((\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n)\, (1 - q)$ if $y_n < (\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}}$. Finally, a Lasso regression on the basic feature set reduces of a few percents the prediction error regarding the voltage magnitude and active power quantities.

III. Optimal Voltage Control

Overvoltages around noon due to large and simultaneous PV injections are expected to be a major contingency in low-voltage grids with high shares of DERs, as shown for the case study in Sec. V-B. Reactive power control is a cost-effective means to regulate the voltage, but in the considered system, reactive power is highly volatile and its forecast is associated with a large uncertainty. In order to show the added value of quantile forecasts, we opt for active power curtailment of the PV injection as measure to keep the voltages below an upper limit, set to 1.05 pu in this study. This simultaneously reduces potential line overloadings. In that respect, this section presents an approach that aims at optimizing the physical DSO position on the intraday market. If an overvoltage can be estimated one hour ahead, the optimal PV curtailments with respect to the estimated state can be computed and we assume that the DSO only compensates the owners of the curtailed PV systems based on the corresponding market price. If an overvoltage is not (fully) eliminated in advance, the DSO has to further curtail PV energy in real time, which is penalized by a higher imbalance price [11]. This results in a trade-off between the risk of curtailing too much energy in advance and facing a higher price for potential adjustments in real time. In this paper, a pure real-time optimization strategy is compared with control strategies based on point and quantile forecasts.

A. Benchmark Real-Time Optimization Strategy (S_B)

In the benchmark strategy S_B, PV curtailment is applied in real time for each time step subject to an overvoltage, i.e.:

$$\max_{k \in \Psi_{\text{B}}} V_k > 1.05, \tag{4}$$

where V_k is the voltage magnitude at bus k, and Ψ_{B} is the set of all buses. PV curtailment is associated with a high imbalance price and the optimization problem is defined as:

$$\min_{P_{n_{\text{PV}}}^{\text{curt}}} \sum_{\Psi_{\text{PV}}} \frac{1}{4} C_{\text{ib}} P_{n_{\text{PV}}}^{\text{curt}}, \tag{5a}$$

subject to

$$0 \leq P_{n_{\text{PV}}}^{\text{curt}} \leq P_{n_{\text{PV}}}^{\text{prod}}, \ \forall n_{\text{PV}} \in \Psi_{\text{PV}}, \tag{5b}$$

$$P_k + \sum_{\Psi_{\text{PV},k}} P_{n_{\text{PV}}}^{\text{curt}} + \sum_{m \in \Omega_k} P_{km} = 0, \ \forall k \in \Psi_{\text{B}}, \tag{5c}$$

$$P_{km} = (V_k)^2 g_{km} - V_k V_m (g_{km} \cos\theta_{km} + b_{km} \sin\theta_{km}), \ \forall k, m \in \Psi_B, \tag{5d}$$

$$Q_k + \sum_{m \in \Omega_k} Q_{km} = 0, \ \forall k \in \Psi_{\text{B}}, \tag{5e}$$

$$Q_{km} = -(V_k)^2 b_{km} + V_k V_m (b_{km} \cos\theta_{km} - g_{km} \sin\theta_{km}), \ \forall k, m \in \Psi_B, \tag{5f}$$

$$0.95 \leq V_k \leq 1.05, \ \forall k \in \Psi_{\text{B}}, \tag{5g}$$

$$V_1 = V_1^{\text{meas}}, \tag{5h}$$

$$\theta_1 = 0, \tag{5i}$$

where $P_{n_{\text{PV}}}^{\text{curt}}$ and $P_{n_{\text{PV}}}^{\text{prod}}$ are the curtailed power and production potential of PV system n_{PV}, respectively. P_k and Q_k are the net active and reactive power consumptions before curtailment at bus k, respectively. P_{km} and Q_{km} are the active and reactive power flows from bus k to bus m, and V_1^{meas} is the voltage magnitude measured at the slack bus. Parameters g_{km} and b_{km} are the line conductance and susceptance from bus k to bus m, θ_{km} is the voltage angle difference between bus k and bus m, and θ_k is the voltage angle at bus k. Ψ_{PV} is the set of all PV systems in the network, $\Psi_{\text{PV},k}$ is the set of all PV systems connected to bus k, Ω_k is the set of all power lines

connected to bus k, and C_{ib} is the imbalance price of curtailed PV energy. The cost function (5a) defines the total curtailment cost per time step of 15 minutes. Constraint (5b) limits the PV power curtailment. The active and reactive node balances as well as the AC power flow equations are defined in (5c)–(5f). Constraint (5g) ensures the voltage boundaries, and constraints (5h) and (5i) set the voltage magnitude and angle references at the slack bus.

B. Optimization Strategy using Point Forecasts (S_{50})

Assuming that point forecasts (i.e., 50%-quantile forecasts) are available, strategy S_{50} is used for each time step where an overvoltage is forecasted hour-ahead:

$$\max_{k\in\Psi_{\mathrm{B}}} \hat{V}_{50,k} > 1.05, \tag{6}$$

where $\hat{V}_{50,k}$ is the point forecast of the voltage magnitude at bus k. Point forecasts of the loads are integrated into the following optimization problem solved at an hour-ahead stage:

$$\min_{P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}}} \sum_{\Psi_{\mathrm{PV}}} \frac{1}{4} C_{\mathrm{m}} P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}}, \tag{7a}$$

subject to

$$0 \le P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}} \le \hat{P}^{\mathrm{prod}}_{n_{\mathrm{PV}}},\ \forall n_{\mathrm{PV}} \in \Psi_{\mathrm{PV}}, \tag{7b}$$

$$\hat{P}_{50,k} + \sum_{\Psi_{\mathrm{PV},k}} P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}} + \sum_{m\in\Omega_k} P_{km} = 0, \forall k \in \Psi_{\mathrm{B}}, \tag{7c}$$

$$P_{km} = (V_k)^2 g_{km} - V_k V_m (g_{km}\cos\theta_{km} + b_{km}\sin\theta_{km})\ \forall k,m \in \Psi_{\mathrm{B}}, \tag{7d}$$

$$\hat{Q}_{50,k} + \sum_{m\in\Omega_k} Q_{km} = 0,\ \forall k \in \Psi_{\mathrm{B}}, \tag{7e}$$

$$Q_{km} = -(V_k)^2 b_{km} + V_k V_m (b_{km}\cos\theta_{km} - g_{km}\sin\theta_{km})\ \forall k,m \in \Psi_{\mathrm{B}}, \tag{7f}$$

$$0.95 \le V_k \le 1.05,\ \forall k \in \Psi_{\mathrm{B}}, \tag{7g}$$

$$V_1 = \hat{V}_{50,1}, \tag{7h}$$

$$\theta_1 = 0, \tag{7i}$$

where $\hat{P}^{\mathrm{prod}}_{n_{\mathrm{PV}}}$ is the point forecast of the potential power production of PV system n_{PV}, and $P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}}$ is the hour-ahead curtailed power of PV system n_{PV}. $\hat{P}_{50,k}$ and $\hat{Q}_{50,k}$ are point forecasts of the active and reactive power consumption at bus k, respectively, and C_{m} is the market price of curtailed PV energy. Note that separate point forecasts based on a simple NN are performed for each PV system output. This optimization problem is similar to problem (5) of the benchmark strategy S_B except that hour-ahead point forecasts replace the realizations and that the associated marginal cost is the market price instead of the imbalance price.

In a second stage, the PV power injection is further adjusted in real time to remove remaining overvoltages. Basically, a similar optimization problem as in strategy S_B is solved whenever condition (4) is satisfied. Nevertheless, the decisions made at the hour-ahead stage must be included by replacing Eqs. (5b) and (5c) with the following constraints:

$$0 \le P^{\mathrm{curt}}_{n_{\mathrm{PV}}} \le P^{\mathrm{prod}}_{n_{\mathrm{PV}}} - P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}},\ \forall n_{\mathrm{PV}} \in \Psi_{\mathrm{PV}}, \tag{8a}$$

$$P_k + \sum_{\Psi_{\mathrm{PV},k}} P^{\mathrm{curtHA}}_{n_{\mathrm{PV}}} + \sum_{\Psi_{\mathrm{PV},k}} P^{\mathrm{curt}}_{n_{\mathrm{PV}}} + \sum_{m\in\Omega_k} P_{km} = 0, \forall k \in \Psi_{\mathrm{B}}. \tag{8b}$$

C. Optimization Strategy using Quantile Forecasts (S_q)

Strategy S_q uses quantile forecasts of the voltage magnitudes to minimize the cost of overvoltages at the hour-ahead stage. For this purpose, we define the uncertainty of the voltage as the difference between the quantile and point forecasts:

$$d\hat{V}_{q,k} = \hat{V}_{q,k} - \hat{V}_{50,k},\ \forall q \in]0,100[,\ \forall k \in \Psi_{\mathrm{B}}, \tag{9}$$

where $d\hat{V}_{q,k}$ and $\hat{V}_{q,k}$ are the voltage uncertainty and forecast for quantile q at bus k, respectively. The curtailment strategy is applied at each time step for which the following condition holds:

$$\max_{k\in\Psi_{\mathrm{B}}} \hat{V}_{q,k} > 1.05. \tag{10}$$

The hour-ahead optimization problem of strategy S_q only differs from problem (7) of S_{50} in the formulation of the voltage limit (7g) and the slack bus voltage reference (7h) which must be replaced by the following constraints:

$$0.95 \le V_k + d\hat{V}_{q,k} \le 1.05,\ \forall k \in \Psi_{\mathrm{B}}, \tag{11a}$$

$$V_1 = \hat{V}_{50,1} - d\hat{V}_{q,1}. \tag{11b}$$

Note that the voltage forecasts are only included in the form of uncertainty of the quantile whereas the resulting voltages are determined by the optimization. In addition, only point forecasts of the power consumption are used in order not to mix the uncertainties from different forecasting sources. Finally, remaining overvoltages detected in real time are handled by the same second-stage optimization as for S_{50}.

IV. Distribution Grid Model

This section presents the distribution network and the associated load, together with the PV and EV active power data, which serves as the model for the evaluation of the forecasting algorithms and control strategies presented in Sec. V. Since this study focuses on data forecasting, special care is taken to create a realistic model based on real data.

A. Grid Topology and Measurements

The system considered for the case study is a weakly meshed 400V low-voltage network fed by a distribution transformer in a residential area of the City of Basel and operated by IWB [12]. It consists of 198 power lines and 196 buses, of which 88 buses are connected to residential loads. About half of the loads are actually measured by smart meters with 15-minute resolution. The yearly energy consumption being known for the other loads, they are assigned power profiles of smart metered consumers with similar consumption located in other residential areas of Basel. Together with the voltage

measurement at the transformer, an observable grid can be achieved. We assume the grid topology to be perfectly known, the loads and PV systems to be voltage independent and the transformer voltage to be maintained after the optimization.

B. Photovoltaic Power Production

In order to simulate a situation where up to 60% of the houses are covered with photovoltaic panels, the power output profiles of 116 PV systems measured by IWB and spread in the entire City of Basel are first selected and normalized by their maximal power value. Second, they are suitably allocated and scaled to the houses based on a tool developed by UVEK [13] and Energie Schweiz [14] which assesses the solar potential of any Swiss rooftop. The solar irradiance measurements are taken from a MeteoSwiss weather station in Basel [15].

C. Electric Vehicle Power Consumption

The consumption profiles of charging EVs are derived from the open data set of the "My Electric Avenue" project, where the driving and charging patterns of more than 200 Nissan Leaf vehicles have been recorded in the United Kingdom over 18 months [16]. After applying data cleaning and filtering, 180 charging profiles at 3.7 kW nominal power are extracted, which corresponds to 30% of the households in the considered grid. Since future home chargers are expected to work between 7.4 kW and 11.1 kW [17], the charging power is scaled up while the charging time is accordingly reduced in order to keep the same energy consumption. EVs associated with a 7.4 kW charger are modeled with the same energy consumption as Nissan Leaf vehicles whereas EVs with a 11.1 kW charger are assumed to have a power consumption close to the Audi e-tron or Tesla models and their energy consumption is multiplied by 1.8 [18]. Finally, charging profiles are allocated to the grid buses according to the algorithm presented in [19] which creates EV clusters. This reflects the social effect of increased willingness to purchase an EV when neighbors also drive EVs.

D. Data aggregation

All aforementioned measurement data are adjusted to a 15-minute resolution over one year. In order to represent different DER penetration levels, multiple data sets are created by adding an increasing number of the EV power consumption and PV production profiles to the initial load profiles. This is summarized in Table III which also indicates the share of houses whose rooftop are equipped with PV panels and the share of households in possession of an EV. In addition, the share of 11.1 kW chargers increases with the DER penetration to simulate a probable decrease of the price gap between 7.4 kW and 11.1 kW chargers. Finally, all bus voltages and power line flows are determined by power flow simulations to complete the system state of the four penetration scenarios.

TABLE III
OVERVIEW OF THE DIFFERENT DER MODIFIED DATA SETS

	Number of PVs (share of houses)	Number of EVs (share of households)	Share of 11.1 kW chargers
DS_{B}	0	0	-
DS_1	39 (20%)	60 (10%)	50%
DS_2	77 (40%)	120 (20%)	62.5%
DS_3	116 (60%)	180 (30%)	75%

V. RESULTS AND DISCUSSION

In this section, we first present the evaluation metrics for the quantile forecasts before commenting on the forecasting results. Finally, the optimization strategies are discussed, for which we use the best forecasting algorithms and feature sets on data set DS_3 that leads to the highest overvoltages.

A. Probabilistic Evaluation Metrics

The Skill Score (SS) is used as the main metric to evaluate the performance of the probabilistic forecasts. It reflects the ability of a quantile forecast to create narrow quantile bands, evaluated by the Normalized Sharpness (NSHRP), while maintaining the reliability of those bands, evaluated by the Average Coverage Error (ACE). The Reliability (REL) represents the percentage of targets that can be captured within the predefined quantile band. These four metrics are defined as follows:

$$\mathrm{REL}(B_Q) = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}_{\hat{y}^{(t)}_{50-Q/2} \leq y_t \leq \hat{y}^{(t)}_{50+Q/2}}, \tag{12}$$

$$\mathrm{ACE}(B_Q) = Q - 100 \cdot REL(B_Q), \tag{13}$$

$$\mathrm{NSHRP}(B_Q) = \frac{100}{T \cdot |\bar{y}|}\sum_{t=1}^{T}\left(\hat{y}^{(t)}_{50+Q/2} - \hat{y}^{(t)}_{50-Q/2}\right), \tag{14}$$

$$\mathrm{SS}(B_Q) = \mathrm{ACE}(B_Q) \cdot \mathrm{NSHRP}(B_Q), \tag{15}$$

where y_t is the target value at time step t, and $\hat{y}^{(t)}_{50-Q/2}$ and $\hat{y}^{(t)}_{50+Q/2}$ are the quantile forecasts defining the lower and upper bounds of the quantile band B_Q at time t, respectively. $|\bar{y}|$ is the mean absolute of the target values, $\mathbb{1}_x$ is the indicator function under condition x, Q is the nominal coverage rate of B_Q, and T is the number of time steps in the test set. A good quantile forecast implies a low sharpness and an ACE close to zero, hence a skill score close to zero.

B. Probabilistic State Forecasting

Figure 2 shows the resulting quantile forecasts of a NN for the active power consumption over three days at a specific bus with PV panels and EVs in the DER penetration scenario DS_3. One can notably see the large PV injection during the first day and the EV consumption in the first two evenings. Most uncertainty appears to come from car charging events such that the additional car charging feature allows for a drastic drop of the forecast uncertainty. The availability of online SM measurements helps to forecast the volatile base household loads and to reduce the uncertainty associated with the PV injection. It also enables the detection of EV charging events with a one-hour time delay if the car charging feature is not provided. The quantile forecasts based on both ML algorithms for the active power flow over a specific line are shown in Fig. 3. Whereas the load of EVs, charging at different time periods, is moderate, the simultaneous power injection of multiple PV

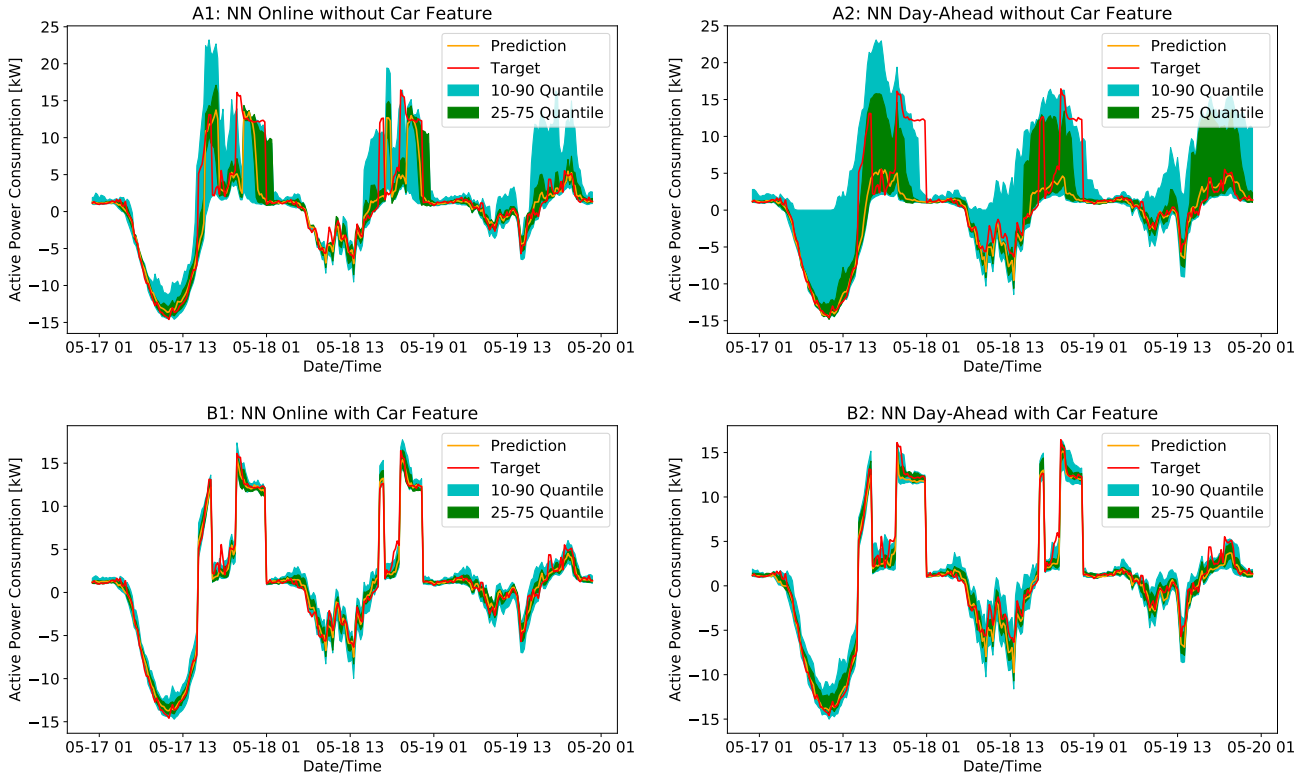

Fig. 2. Active power consumption forecasts at a specific bus with PV panels and EVs for DS_3. Subfig. A1: Neural Network (NN) with online SM measurements. Subfig. A2: NN with only day-ahead SM measurements. Subfig. B1: NN with online SM measurements and additional car charging feature. Subfig. B2: NN with only day-ahead SM measurements and additional car charging feature.

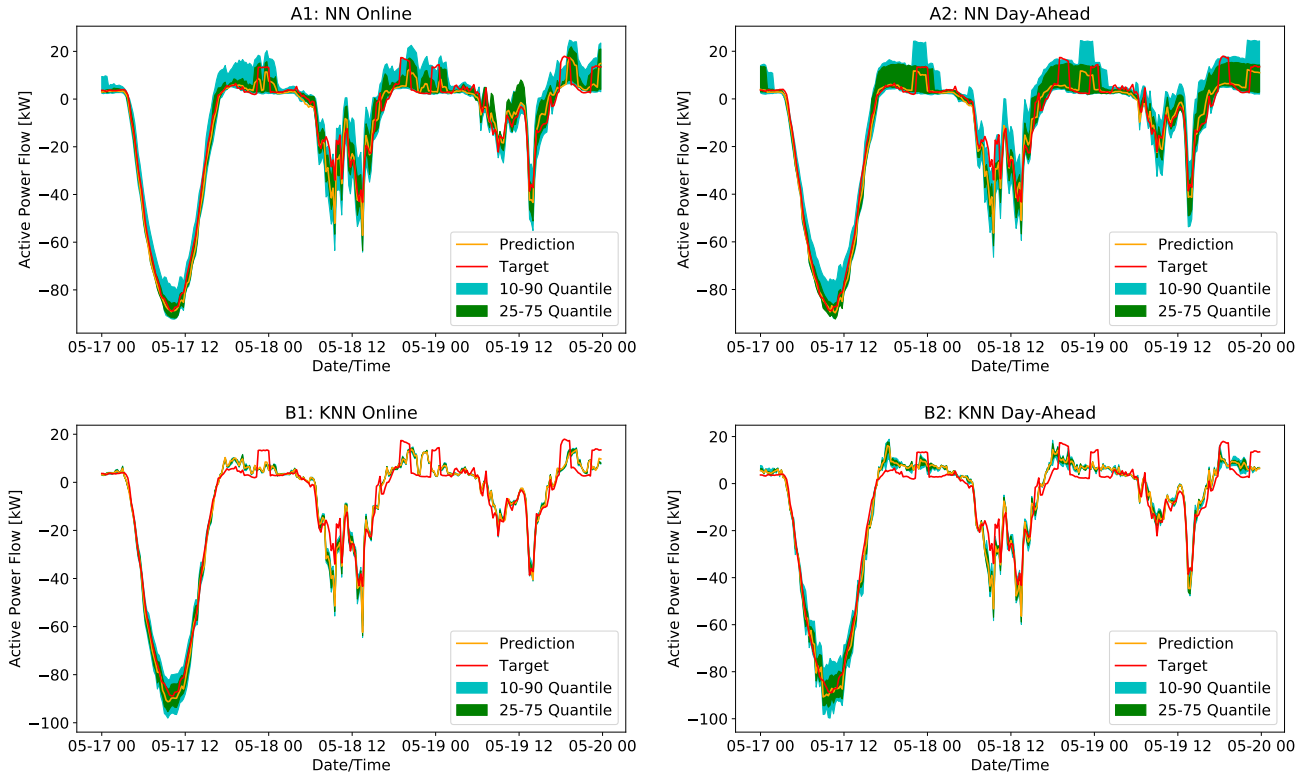

Fig. 3. Active power flow forecasts at a specific line for DS_3. Subfig. A1: NN with online SM measurements. Subfig. A2: NN with only day-ahead SM measurements. Subfig. B1: KNN with online SM measurements. Subfig. B2: KNN with only day-ahead SM measurements.

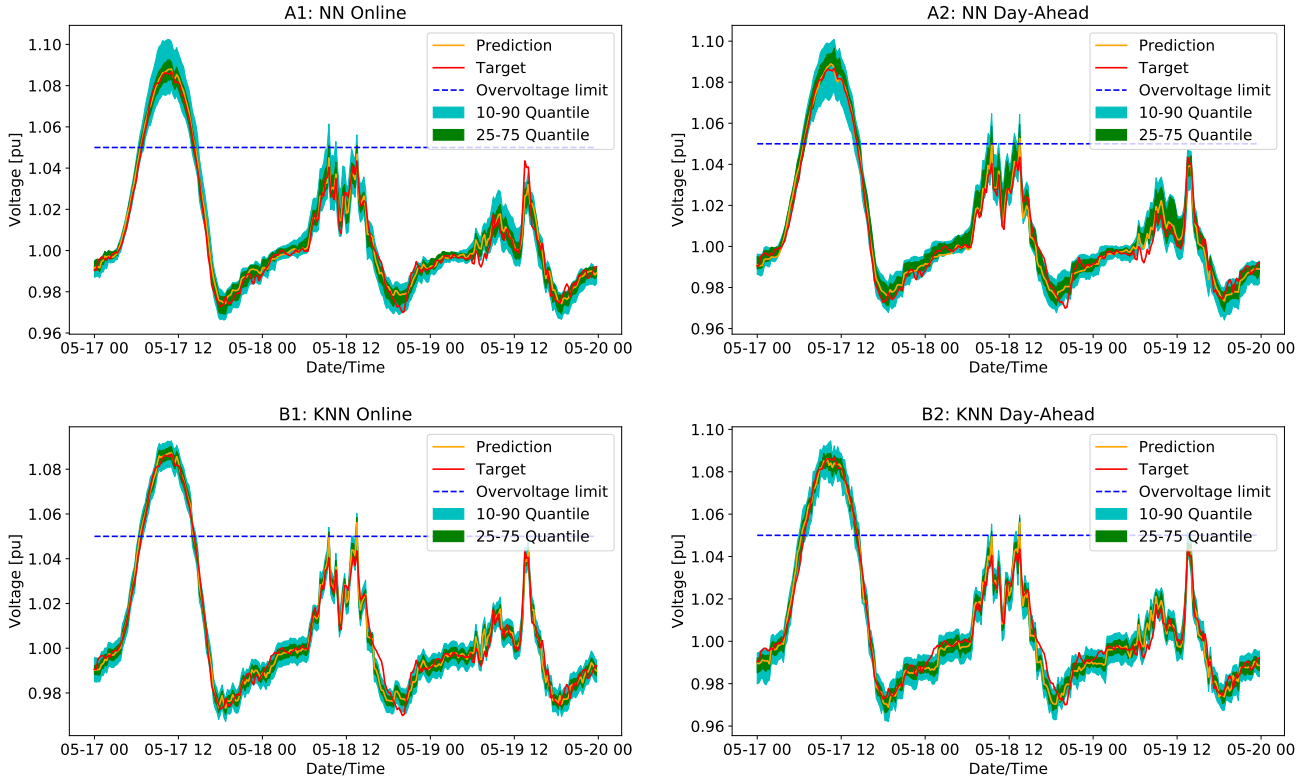

Fig. 4. Voltage magnitude forecasts at a specific bus with PV panels and EVs for DS_3. Subfig. A1: NN with online SM measurements. Subfig. A2: NN with only day-ahead SM measurements. Subfig. B1: KNN with online SM measurements. Subfig. B2: KNN with only day-ahead SM measurements.

systems is clearly visible. The NN algorithm can reasonably forecast the power flow and properly detect the periods with higher uncertainty, even without online SM measurements. However, the KNN algorithm produces very narrow quantile bands which fail to encompass the target values, except during high PV injection time. This phenomenon is observed for each forecast where the difference between the minimum and maximum target values is relatively large. The reason lies in the nature of KNN where the quantile predictions are linear combinations of previous target values. Hence, the bounds of the quantile bands are the results of problem (3) where the optimal weights are applied for all time steps such that the bands get narrower when the target values get closer to zero.

Concerning voltage magnitude forecasts, Fig. 4 compares the outcome of the NN and KNN algorithms with and without online SM measurements. All variants are quite accurate. In this scenario with the highest DER penetration, the shape of the voltage curve is barely impacted by the EV load. In contrast, during sunny days, voltage values largely exceed the overvoltage limit whereas on cloudy days, the detection of an overvoltage depends on the considered quantile. In this case, the accuracy of the quantile forecasts is determinant when used in the control strategies. While still being reliable, KNN tends to produce narrower quantile bands than the NN. Reactive power consumptions and reactive power flows are not explicitly shown since they are extremely volatile and hence hardly predictable in any case.

Figures 5 and 6 compare the performance of the NN and KNN algorithms in the form of box and whisker plots, where each data point represents the skill score for a single line power flow and bus voltage, respectively. The central bar indicates the median value, the small red square is the mean value, the box corresponds to the Interquartile Range (IQR) and the ends of the whiskers define $1.5 \times \mathrm{IQR}$ below and above the lower and upper quartiles, respectively. Figure 5 indicates that the NN outperforms KNN to forecast active power flows in all considered cases. This is explained by the specifically poor reliability of KNN which generally creates too narrow quantile bands that miss the target values, as illustrated in Fig. 3. In addition, the ACE of KNN algorithm is proportionally increasing with the uncertainty (i.e., growing DER penetration). The car charging feature and the availability of online SM measurements are beneficial to a lesser extend. Similar outcomes can be observed for active power consumption forecasts. In contrast, Fig. 6 shows the better skill score for KNN compared to the NN, due to a better reliability and lower sharpness. Indeed, the narrower bands produced by KNN still properly envelop the target values in the forecast of voltage magnitudes since all values lie in the same range (i.e., around 1 pu). Being based on the linear combination of similar grid states, KNN is therefore well suited to predict voltages. In terms of further features, only online SM measurements are beneficial for voltage forecasts.

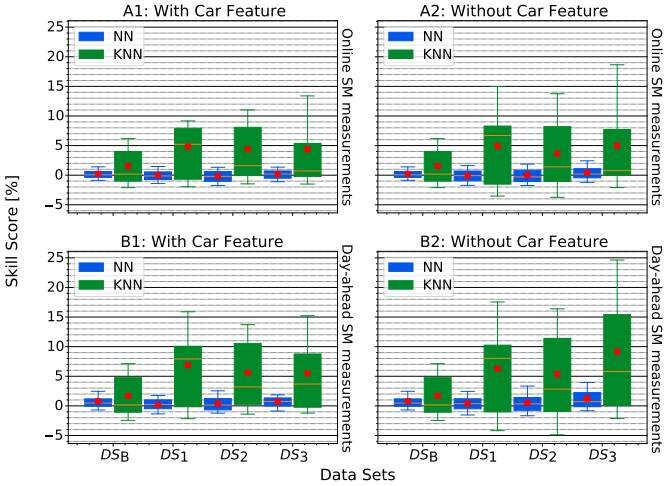

Fig. 5. Skill score for B_{80} of the active power flow for all lines and time steps in the test set. Each subfigure refers to a different feature set related to the availability of online SM measurements and of the car charging feature.

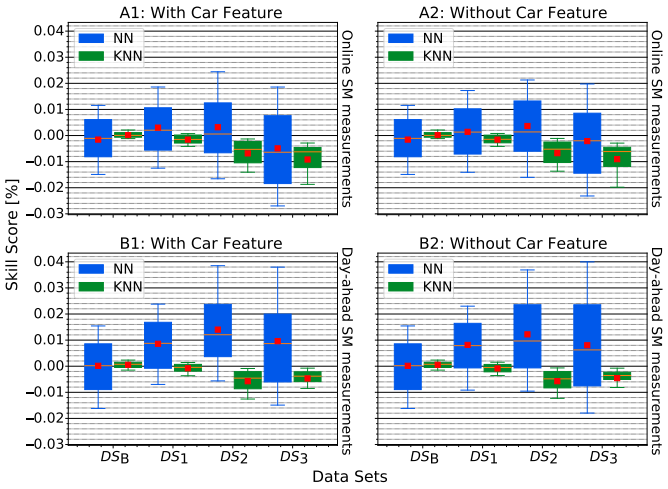

Fig. 6. Skill score for B_{80} of the voltage magnitude for all buses and time steps in the test set. Each subfigure refers to a different feature set related to the availability of online SM measurements and of the car charging feature.

C. Optimal Voltage Control

Having shown the best performance in terms of voltage forecasts, the KNN algorithm with online SM measurements is used. However, no car charging feature applies here. For the power consumptions, the NN algorithm with online SM measurements and car charging feature is considered. Power line flows are not directly used since they are implicitly given by the OPF problems. In addition, on average an ACE $\approx -6\%$ is observed for the 50%-quantile voltage forecasts at time steps where an overvoltage is predicted, which indicates that the point forecasts tend to overestimate more than underestimate the voltage at these time steps. This gives the incentive to relax the voltage limit to the 44%-quantile in strategy S_q. Alternatively, due to a cheaper market price than imbalance price, the DSO might want to remove as much overvoltage as possible at the first stage and even accept superfluous PV curtailment. This would justify the use of a higher quantile in strategy S_q (e.g., 62.5%-quantile in this case study). Moreover, since the results are evaluated on a Swiss grid, the market price C_m for electrical energy is set to 40 €/MWh which roughly corresponds to the average Swiss spot market price [20]. Subsequently, the Swiss imbalance price is defined as [11]:

$$C_{\mathrm{ib}} = 1.1 \cdot (1.2 \cdot C_{\mathrm{m}} + p_{\mathrm{ib}}), \tag{16}$$

where C_{m} and C_{ib} are the market and imbalance prices for the curtailed PV energy in €/MWh, respectively, and p_{ib} is an imbalance penalty equal to 10 €/MWh. Considering the increasing share of volatile DERs in future distribution networks, the imbalances are expected to increase, which could lead to an increase of the imbalance price with respect to the market price. Hence, we also consider a future situation where the imbalance price would be doubled.

TABLE IV
OPTIMIZATION RESULTS FOR DIFFERENT STRATEGIES IN CURRENT AND POTENTIALLY FUTURE PRICE SITUATIONS

	S_B	S_{50}	S_{44}	$S_{62.5}$
HA power curtailment [MWh]	0.0	22.3	21.6	23.7
RT power curtailment [MWh]	22.3	2.1	2.4	1.5
Total power curtailment [MWh]	22.3	24.4	24	25.2
Total cost in current situation [€]	1422	1022	1017	1044
Total cost in future situation [€]	2843	1154	1172	1142

Table IV compares the benchmark and forecast based strategies for both price situations. The total cost consists of the cost of the Hour-Ahead (HA) and of the Real-Time (RT) curtailments for a period of 10 weeks evenly distributed over a year, as shown in Fig. 1. Although the forecast based strategies curtail in total more PV power than the pure real-time optimization because of the prediction errors, they lead to a clear reduction of the total cost (i.e., about 27%). In the current price situation, S_{44} is the most cost-efficient strategy. Since the imbalance price is only about 50% higher than the market price and the point forecast tends to overestimate the voltage in overvoltage situations, it is preferable to enable some more remaining overvoltages that are handled in real time. Conversely, if the imbalance price increases, it gets profitable to accept more and even too much curtailment one hour ahead, as shown by the lowest total cost for strategy $S_{62.5}$. Note that the actual value of the market and imbalance prices influences only the total cost, not the optimal curtailment.

Finally, based on strategy S_{44} at a selected time step, Fig. 7 gives insight into the intermediate voltage predictions in the 50%-quantile before and after the hour-ahead optimization and into the final voltage realizations for all buses in the network. Since this strategy reduces the overvoltages only to the 44%-quantile, the upper subplot still shows some remaining overvoltages predicted for the 50%-quantile at a few buses. Nevertheless, at this time step, the lower subplot indicates that all the voltages are actually well below the limit already after the hour-ahead stage. This means that no real-time adjustment is required. Due to the tendency of the point forecast to overestimate overvoltages, strategy S_{44} profits of exactly such

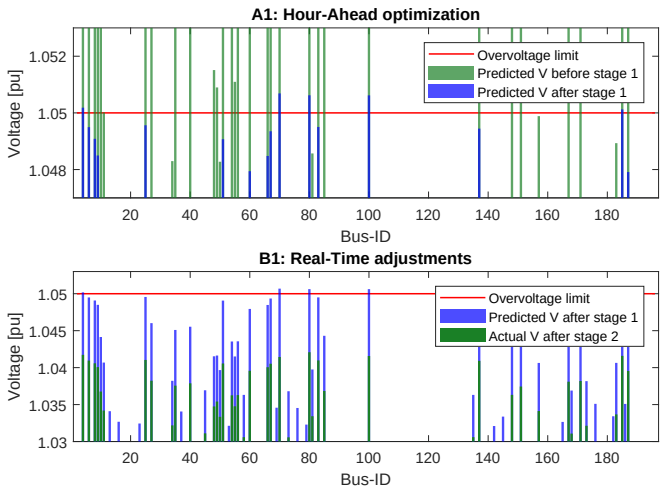

Fig. 7. Intermediate voltage point forecasts and final voltage realizations resulting from strategy S_{44} at a selected time step

scenarios where the overvoltage can be completely eliminated at the hour-ahead stage by curtailing less energy than the point forecast would have predicted.

VI. Conclusion

To sum up, this paper is meant to encourage the use of quantile forecasts in OPF schemes. Hence, we present a comprehensive approach for hour-ahead probabilistic state estimation whose quantile predictions are subsequently used in a control scheme that optimally reduces the costs associated with overvoltages. For that purpose, we design a quantile NN that can accurately forecast power quantities and properly quantify the corresponding uncertainty, and we propose a novel quantile KNN algorithm that outperforms the NN for the prediction of voltage magnitudes. The hour-ahead forecast evaluation is performed on a real LV grid with various DER penetrations. It highlights the particularly large uncertainty on bus power consumptions and line power flows caused by EVs and the added value of knowing their charging starting time and duration to substantially increase the forecast reliability. The availability of online SM measurements is only marginally beneficial. Moreover, the increasing DER penetration barely affects the accuracy of the algorithms when evaluated by the skill score which combines the average coverage error and the normalized sharpness of quantile bands. In addition, we suggest a way to integrate quantile forecasts in an optimal voltage control scheme where PV power curtailment can be decided hour-ahead at a lower cost. While the general use of forecasts in such a scheme is definitely worthwhile, the paper shows that quantile forecasts can reduce the costs even further compared to point forecasts. Nevertheless, the exact quantiles to consider and the corresponding cost reduction depend on the price difference between hour-ahead and real-time curtailment.

Based on these promising outcomes, a next step is to investigate the optimal quantiles that maximize the cost reduction depending on the price situation. Different quantiles could also be integrated in the same optimization problem and even directly contribute to the cost function instead of acting on the voltage limit constraint. The performance of the suggested control scheme should also be compared with other probabilistic approaches such as stochastic or chance-constrained optimization where the grid quantities are seen as random variables with a certain probability distribution. In addition, multiple means to control voltages such as reactive power control, online tap changing of transformers, and the use of the EV flexibility should be considered, preferably in a three-phase system due to the unbalanced nature of power flows in distribution grids. Finally, the presented OPF scheme could be adapted to congestion management, demand-side management, or any application subject to uncertainty that can benefit from short-term forecasts.

References

[1] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875–3883, 2016.

[2] Y. Weng, R. Negi, C. Faloutsos, and M. D. Ilić, "Robust data-driven state estimation for smart grid," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1956–1967, 2016.

[3] B. P. Hayes, J. K. Gruber, and M. Prodanovic, "A closed-loop state estimation tool for MV network monitoring and operation," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 2116–2125, 2014.

[4] R. Bessa, G. Sampaio, V. Miranda, and J. Pereira, "Probabilistic low-voltage state estimation using analog-search techniques," in *IEEE Power Systems Computation Conference (PSCC)*, 2018.

[5] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[6] G. Dahua, W. Yi, Y. Shuo, and K. Chongqing, "Embedding based quantile regression neural network for probabilistic load forecasting," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 2, pp. 244–254, 2018.

[7] Y. Wang, D. Gan, M. Sun, N. Zhang, Z. Lu, and C. Kang, "Probabilistic individual load forecasting using pinball loss guided lstm," *Applied Energy*, vol. 235, pp. 10–20, 2019.

[8] S. Karagiannopoulos, L. Roald, P. Aristidou, and G. Hug, "Operational planning of active distribution grids under uncertainty," in *10th Bulk Power System Dynamics and Control Symposium (IREP)*, 2017.

[9] L. Roald and G. Andersson, "Chance-constrained AC optimal power flow: Reformulations and efficient algorithms," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2906–2918, 2017.

[10] T. Zufferey, A. Lepouze, and G. Hug, "Inadequacy of standard algorithms and metrics for short-term load forecasts in low-voltage grids," in *IEEE PowerTech*, 2019.

[11] G. Koeppel and D. Reichelt, "Power Market I," *Lecture, ETH Zurich*, 2019.

[12] IWB. Accessed on 2020-03-11. [Online]. Available: www.iwb.ch

[13] UVEK. Wie viel Strom oder Wärme kann mein Dach produzieren? Accessed on 2020-03-11. [Online]. Available: www.uvek-gis.admin.ch/BFE/sonnendach

[14] Energie Schweiz. Solarrechner. Accessed on 2020-03-11. [Online]. Available: www.energieschweiz.ch/page/de-ch/solarrechner

[15] MeteoSwiss. IDAWEB. Accessed on 2020-03-11. [Online]. Available: gate.meteoswiss.ch/idaweb

[16] J. Cross and R. Hartshorn, "My Electric Avenue: Integrating electric vehicles into the electrical networks," 2016.

[17] Bosch Automotive Service Solutions. Electric vehicle solutions. Accessed on 2020-03-11. [Online]. Available: www.boschevsolutions.com/charging-stations

[18] F. Ferrando, "Potential impact of electric vehicles on the swiss energy system," *Master Thesis, ETH Zurich*, 2019.

[19] J. Stiasny, "Sensitivity analysis of EV impact on distribution grids based on Monte-Carlo simulations," *Master Thesis, ETH Zurich*, 2019.

[20] European Energy Spot Market. EPEX SPOT. Accessed on 2020-03-11. [Online]. Available: www.epexspot.com