IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 February 2023, accepted 15 February 2023, date of publication 27 February 2023, date of current version 2 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3249459

RESEARCH ARTICLE

Predicting Renewable Curtailment in Distribution Grids Using Neural Networks

ELENA MEMMEL, THOMAS STEENS, SUNKE SCHLÜTERS, RASMUS VÖLKER, FRANK SCHULDT, AND KARSTEN VON MAYDELL

DLR Institute of Networked Energy Systems, 26129 Oldenburg, Germany

Corresponding author: Elena Memmel (elena.memmel@dlr.de)

ABSTRACT The growing integration of renewable energies into electricity grids leads to an increase of grid congestions. One countermeasure is the curtailment of renewable energies, which has the disadvantage of wasting energy. Forecasting congestion provides valuable information for grid operators to prepare and instruct countermeasures to reduce these energy losses. This paper presents a novel approach for congestion prediction in distribution grids (i.e. up to 110 kV) considering the n-1 security criterion. For this, our method considers node injections and power flow and combines three artificial neural network models. The analysis of study results shows that the implemented neural networks within the presented approach perform better than naive forecasts models. In the case of vertical power flow, the artificial neural networks also show better results than comparable parametric models: average values of the mean absolute errors relative to the parametric models range from 0.89 to 0.21. A high level of accuracy can be achieved for the neural network that predicts the loading of grid components with a F1 score of 0.92. Further, also with a F1 score of 0.92, this model shows higher accuracy for the distribution grid components than for those of the transmission grid, which achieve a F1 score of 0.84. The presented approaches show good potential to support grid operators in congestion management.

INDEX TERMS Power system operation, distribution grid, congestion management, renewable power curtailment, artificial neural network, short-term prediction, vertical power flow.

NOMENCLATURE

ANN	artificial neural network.
CHP	combined heat and power unit.
DSO	distribution system operator.
EHV	extra high voltage.
HV	high voltage.
LV	low voltage.
LSTM	long short-term memory.
MAE	mean absolute error.
ML	machine learning.
MSE	mean squared error.
MV	medium voltage.
PV	photo voltaic.
RE	renewable energies.
RMAE	relative mean absolute error.
RMSE	root mean square error.
SCADA	supervisory control and data acquisition.
TPE	tree-structured parzen estimator.
TSO	transmission system operator.
WP	wind power.
WTPC	wind turbine power curve.

The associate editor coordinating the review of this manuscript and approving it for publication was Emilio Barocio.

I. INTRODUCTION

Due to energy transition and the consequential rise of renewable energies (RE) integrated in electricity grids, the amount of congested grid components have increased [1]. One way to ensure save operation within the limits and counteract overload is to curtail RE [2]. In Germany the amount of RE curtailment caused by grid congestions increased from 3.74 TWh to 6.15 TWh [3]. As curtailment results in a loss of energy, congestion management needs to be optimized [4].

Forecasting upcoming congestion can give network operators additional information and time to plan countermeasures and ultimately decrease curtailment. This can contribute to optimization of congestion management [5]. Furthermore in Germany the regulation regarding curtailing RE is changed from former infeed management (*Einspeisemangement* defined by §14 EEG), where curtailment is carried out on demand by grid operator [6], to ''Redispatch 2.0'' in October 2021 [7], where RE curtailment is now regulated by §§13,13a, 14 EnWG. As a consequence predictions of grid congestions are also obligatory for distribution grids [8].

So forecasts for distribution grids are required, but due to unpredictable power flow challenging to determine [9]. History shows that deep learning is capable solving complex systems with increasing accuracy and provides the advantage of being flexible [10]. Further, Ilić et al. stated that artificial neural networks (ANNs) have the advantage to use a ''reduced input space'' and are suitable to be applied in situations that require lots of parallel forecasts [11]. And Staudt et al. showed that an ANN is a promising tool for determining redispatch [12]. Thus, the idea of this paper is to use ANNs for predictions in the distribution grid. Since grid operation require a high level of security, the presented assessment of the model's accuracy is a first step towards evaluating the integration of ANNs into operative grid management. As parametric models already exist that provide the required predictions [13], [14], a comparison of the ANNs with the parametric approaches is also conducted. Because measured transformer power are recorded in SCADA (supervisory control and data acquisition) systems, past grid injections are available and former grid congestions can be reconstructed (as shown in [14]). This data can then be utilized to predict upcoming curtailments in distribution grids using ANN.

In literature ANN were already successfully applied in the context of congestion detection [5], RE power [15], [16], [17] and consumption prediction [18]. An overview of those ANNs is given in Table 1. Staudt et al. proposed an ANN for day-ahead prediction of redispatch of individual power plants by using empirical data from the German electricity market [12]. Line congestions in transmission grids are predicted via ANN in [5]. Srivastava et al. proposed an approach relying on Monte Carlo simulation and probabilistic load flow to predict congestions occuring in LV grid levels. In their study generation and consumption are modeled by applying meteorological data gathered by fisheye lens cameras to predict PV power production and ANNs for load forecast [18]. Another aproach is to use ANN for congestion prediction in electricity grids. Fainti et al. used an ANN trained by the Levenberg-Marqardt algorithm with implemented Bayesian regularization in order to predict congestion on each of the three phases of a power distribution line [19]. Alali et al. trained two ANNs to get the probability of a congested line in the first model (using complex bus voltage, bus active- and reactive power) and the source of congestion (the causing bus) in the second one [20]. The model was trained for a modified 4.16 kV IEEE standard 12-bus distribution grid.

In order to predict possible congestions occuring in the distribution grid, injections into grid also have to be predicted. The injections can be simulated by modeling RE injection and load separately [18], using meteorological information and information about consumption for phase prediction [19]. Focusing on congestion prediction in [20] they used active and reactive power of each bus as input features.

Following requirements for congestion prediction are addressed by the proposed models of this study: Detailed information of node-injections into the 110 kV grid level [21], [22], consideration of n-1 security (which means that one grid component can fail and grid security is still maintained) [23], [24], and the amount of required RE curtailment should become determinable. As it is aimed to develop a method applicable for grid operation, it should be validated against actual grid data. The paper makes the following contribution:

- The need for curtailment forecasts is addressed, by presenting a model that detects congestions and quantifies overloading to determine the power to be curtailed.
- The methods presented are ANNs, that predict vertical power, for MV/HV transformers and wind farms connected to HV grid level, and congestions under consideration of n-1 security. These methods address the following needs identified in the context of congestion occuring in distribution systems:
  - Time series prediction of vertical power via ANN, considering individual injection characteristic with a time step of 15 minutes.
  - Component loading prediction for 110 kV distribution grid via ANN, considering n-1 security, which functionality includes detecting congestions and so enables the determination of the power that needs to be curtailed.
- ANN approaches for vertical power prediction and congestion forecast are compared to a parametric model using same input data. These are contrasted and evaluated to give an outlook on the potentials of both approaches.

The remainder of the paper is structured as follows: In Section II the methodology of different ANNs for predicting vertical power on MV/HV transformers, generated power of wind farms injected at HV transformers and prediction of component loading considering contingencies are presented. In Section III the accuracy of the models are analyzed and compared to different approaches. The key findings are discussed in Section IV and finally concluded in Section V.

II. METHODOLOGY

The scope of this study is to enable prediction of curtailment by designing a machine learning (ML) forecast model for congestions induced by RE in distribution grids.

To achieve this objective three ANNs are designed. Fig. 1 shows the interrelation between the models and the components (each grid component type is assigned a number)

TABLE 1. Overview of ANNs presented in literature for the prediction of wind power, vertical power flow and grid congestions.

Publication	Applied Model	Forecast Horizon	Short Description	Input data	Comparison
Braun et al. [9]	LSTM implemented daily retraining	15 minutes time step Forecast horzion: between 1-48 hours	Vertical power prediction	Weather data (sun position, directions of wind speed, temperature forecast albedo and surface solar daiation downwards, surface pressure, total precipitation), measurements of vertical power	see Table 9
Catalao et al., [15]	ANN (feed forward; 3 layerd) trained by Levenberg Marquard algorithm	Forecast horizon: 1 day Timestep: 15minutes	Short-term wind power prediction in Portugal	historical wind power data	not comparable: wind power prediction for portugal
Bhaskar et al. [16]	I)Adaptive wavelet neural network II)ANN	trained for hour-ahead forecasting Forecasted output is recursively used for prediction up to 30 h ahead	Wind power forecasting	Wavelet layer Hourly averaged 10 min wind speed, wind power	not comparable: takes two stage forecasting
Lin et al. [17]	ANN (5 layered) Weigth initialization: algorithm of xavier activation function: RELU cost function: MSE	measurements per second data from 07.2018-06.2019	prediction of wind power generation of one turbine based on SCADA data base	wind speed measurements of 4 different heights, nacelle orientation, yaw error, average+measured blade pitch, blade pitch angle of 3 blades, ambient temperature	not comparable: model predicts the output of one turbine
Srivastava et al. [18]	ML techniques using sky images ANN probabilistic power flow	PV production forecast:1 minute ahead load forecast: 1 hour ahead congestion forecast horizon: 1 minute - 24 hours ahaed	support tool for DSOs for congestion forecasts	network parameters(branch ampacity, transformer rated capacity), ANN-load: temperature, hour of day, day of week, type of day, monthly growth rate(load), power(of load)	not comparable:considers only MV and LV grid components focus on PV injections
Staudt et al. [12]	ANN (4 layered) Activation function of output layer: sigmoid function 10 folded cross-validation	forecast horizon: day ahead	Prediction of redispatch in German electricity market	day ahead load forecast,-generation forecast, -price forecast, -solar and wind forecast, weekday, season	not comparable: focus on redispatch
Staudt et al. [5]	ANN error minimization: Adam activation function: RELU	hourly time step considered period:01.2015-05.2017	Prediction of transmission line congestions	Congestions data of German TSO, day ahead forecast of net input, -load forecast, -scheduled generation forecast, -spot price forecast, -solar and wind forecast, -forecast of import and export, weekday	see Table12
Fainti et al. [19]	ANN (3-layered) Training: Levenberg-Marquadt Bayesian regularization: to avoid overfitting	measuremetns rate: every 15 minute training data: first two weeks of every month of 2000	Prediction of congestions for all three-phases in a distribution system	wind magnitude, solar radiation, temperature, time, overall demand	not comparable: no realistic grid (used grid:IEEE grid)
Alali [20]	Two ANNs - both 3-layered Bayesian regularization Training: Levenberg-Marquardt cost-function: MSE	considers: 24 hour scenario	Line congestion forecast in active distribution systems	First ANN: Active and reactive power of each bus Second ANN: output of first ANN: possible congestions	not comparable: used IEEE- 13 bus grid - no realistic grid data

of the 110 kV distribution grid (the grid topology was first presented in [14]). To identify congestions, injections into the grid must be known. So, the vertical power on MV/HV transformers is predicted via an ANN, which is designed using meteorological forecasts and measurements of transformer power. The consumed and generated power of the corresponding transformers are represented by bus bar 2 and 4 in Fig. 1. The vertical power flow on transformers connecting wind parks directly to the 110 kV grid (represented by 3), is predicted by a second ANN.

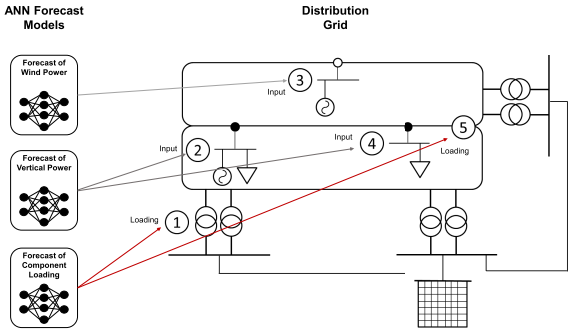

FIGURE 1. Description of how the ANNs forecasts are connected to the 110 kV distribution grid. Scheme of grid topology was first presented in [14].

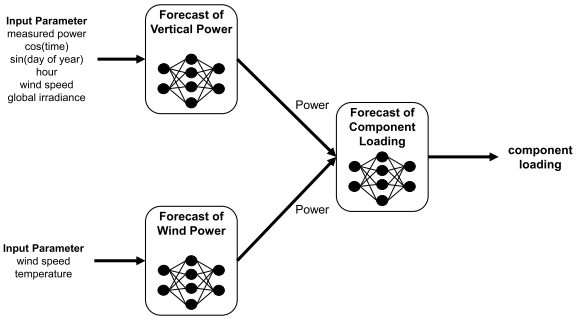

FIGURE 2. Possible combinations of ANNs leading to congestion prediction.

Knowing the power injections into the grid enables the next step: The determination of possible overloadings using the third ANN presented in this paper. Congestions must be determined taking into account n-1 security [23], [24]. Therefore, a contingency analysis has to be considered. The third ANN predicts the maximum load resulting from a contingency analysis for grid components, like lines of the 110 kV grid (described by 5) and HV/EHV transformers (symbolized by 1).

To summarize the proposed models of this paper: Three ANNs are presented, which contribute separately to predict congestions. The ANN for vertical power prediction enables DSOs of the MV grid level to predict congestions occurring on their transformers. Furthermore, the ANNs for vertical power prediction and wind power (WP) predictions provide time series of power infeed, which is required as input to calculate load flows of the distribution grid. And the ANN for forecasts of grid component loading enables determinations of congestions considering n-1 security.

Combining these three models, as visualized in Fig. 2, enables a holistic approach to predict RE caused congestions, which is required to determine RE curtailment. The injections of the nodes are predicted by the ANN for vertical power and the ANN for WP. These node in-feeds are input parameters to the ANN for predicting grid component loadings, thus enabling congestion detection. The advantage of this combination is that it contributes to the knowledge of the whole system, i.e. to the detection of possible congestions on MV/HV transformers and on 110 kV lines. However, in this paper the models are considered separately.

One drawback of ANNs are that they are black boxes. If they do not fulfill the security requirements of the grid operator, the presented approach has the advantage that those separate models can be exchanged with parametric methods. So, a comparison between ANN models and parametric approaches is described in Section III-E of this paper. Another advantage of this approach, is that it allows an evaluation of the intermediate results.

A. ANN SETUP

With regard to redispatch forecasts with a forecast horizon of 3 hours up to 33.5 hours (as day ahead forecast is needed at 2:30 p.m the day before) are required. Therefore, different forecast periods are considered. For all three ANN models, a forecast horizon of 12 steps, i.e. three hours, was chosen. In case of the ANN for vertical power prediction a forecast horizon of 60, i.e. 15 hours, is additionally used. The separation of data and the supervised training of each ANN is structured as follows: The data set, which consist of associated input and output values [25], is divided into a set for training and one for testing.

The training period for the ANN of WP prediction was 01.01.2015 - 31.12.2015 and the testing period was 01.03.2016 - 16.12.2016. For the vertical power ANNs data of the years 2015 and 2016 were available and for the ANN of the component loading the data period 25.03.2016 - 16.12.2016 was used. These ANNs had been trained considering cross-validation [26], by splitting the data into 5 sets. As the data are time series, the temporal order of each set is maintained. In this study the ANNs are trained once.

The presented ANNs consist of multiple layers, which are called dense layers in case of feed-forward layers (as described in [27]) and Long Short-Term Memory (LSTM) layers, where information is stored by "gates" [28]. To avoid overfitting, early stopping (described in Section 4.3 of [27]) is implemented in each ANN.

The output of a neuron from a feed forward layer is defined in [27] by

$$x_i = f(\xi) \tag{1}$$

where ξ denotes the potential of the i_{th} neuron and is defined by

$$\xi = \vartheta_i \sum_{j \in \Gamma^{-1}} \omega_{ij} x_j \tag{2}$$

Γ denotes a function assigning a subset containing the output of connected neurons from previous layer $\Gamma(i)^{-1} \subseteq V$ to each neuron i, ω_{ij} denotes the weight of the connection between the neurons i and j, and ϑ_i is the threshold coefficient of the i_{th} neuron. For further information see [27].

In case of LSTMs the neurons consists of a memory cell. The output $y^{c_j}(t)$ of the j-th cell c_j at time t is defined in [28] by

$$y^{c_j}(t) = y^{\text{out}_j}(t) h(s_{c_j}(t)) \tag{3}$$

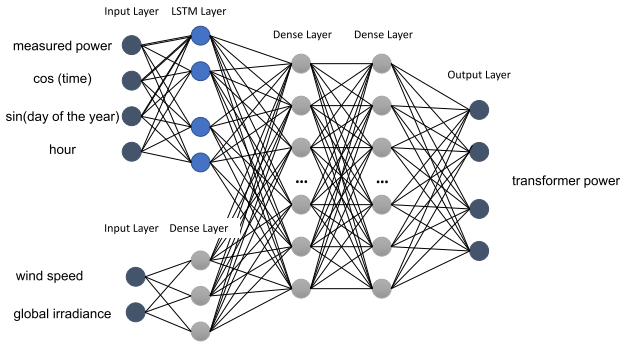

FIGURE 3. ANN architecture for MVHV power prediction.

where y^{out_j} describes the "output gate", a function which prevents perturbation of irrelevant information of following units. h describes the function, that scales the output of the cell, and s_{c_j} defines the internal state of the cell.

For further information regarding LSTM see [28]. For optimization the "adam" algorithm is used to update the weights, which has the advantage of a adapting learning rate. [29]

B. PREDICTION OF VERTICAL POWER FLOW ON MV/HV TRANSFORMERS

Vertical power flow on MV/HV transformers is an aggregation of the generated RE power and the connected consumption. The transformers differ in their amount of installed RE units, or connected loads. So each shows a different characteristic regarding power backfeeded into the HV level and power injected into the MV level. These characteristics can be clustered into five groups based on their ratio of infeed to backfeed, as proposed in our earlier study [13] (See Figure 11, 12 in [13]). These clusters are named according to their dominant power flow direction. In the following the power flow from HV into MV grid level is called injection, and the power flow from MV to HV grid level is called backfeed. Accordingly the resulting transformer clusters which have more frequent backfeeds into the HV grid level, are called *mainly backfeed*, and *backfeed*. Those that show more fequent injections to the MV grid level are referred to as *mainly demand*, and *demand* since in these cases demand exceeds generated power.

The available REs in the simulated region are WP, photovoltaic (PV) and combined heat and power (CHP) units. Accordingly the descriptive features chosen are wind speed, global irradiance, historic measured vertical power, and time.

The corresponding architecture consists of two input layers, one for the historic features and one for the meteorological forecasts. The historic input layer is connected with a LSTM layer, while the meteorological input layer is connected with a dense layer. Both, the LSTM and the dense layers are concatenated and connected to two additional dense layers.

For each transformer in the model region the number of nodes for each layer of the ANN is individually adjusted by a hyperparameter optimization to ensure optimal results. The optimization is performed by the python package optuna [30].

TABLE 2. Hyperparameters for ANN predicting Transformer Power.

Parameter	Selected Values
Historic-/Meteorological Input Layer	4/2
Output	length of forecast
Training algorithm function	Adam [29]
Hidden Layer 1 - LSTM - function	Tangens hyperbolicus
Hidden Layer 1 - Dense - function	Relu activation function
Hidden Layer 2 - Dense - function	Tangens hyperbolicus
Hidden Layer 3 - Dense - function	Linear activation function
Hidden Layer 1 - LSTM - size	32 - 128
Hidden Layer 1 - Dense - size	16 - 112
Hidden Layer 2 - Dense - size	64 -128
Hidden Layer 3 - Dense - size	64 - 128
Weight update method	batch-mode
Early Stopping	patience of 4
Learning rate	1E-3
batchsize	150
Loss function	MAE

Optuna was configured to minimize the MAE of the resulting predictions using "Tree-structured parzen estimator" (TPE) sampling algorithm. For more information regarding TPE see [31].

The parameters for the number of nodes for each layer of the optuna optimization model has been set to discrete steps with stepsize of 8, a minimum number of 8 and a maximum number of 128. The nodes per layer are not mapped individually per transformer, but the range from the smallest to the largest number of occurring nodes is given. The occuring hyper parameters are shown in Table 2.

The derived networks are then trained individually for every transformer.

C. PREDICTION OF WIND POWER GENERATION

Some wind farms are directly connected to the 110 kV distribution grid via a transformer. Because these transformers have no consumers associated an additional ANN-architecture is required.

As for these wind farms the installed capacity is the only information available, an adjusted wind turbine power curve is used to calculate generated wind power from wind speed model data (for further information regarding the data see Section III-A). A more detailed description of the calculation of WP can be found in Section II-B.1 of our previous paper [14]. The resulting power is used as target feature for the following ANN.

In literature approaches using ANN to predict WP generation have been proposed [15], [16], [17]. As the corresponding architectures have one to three hidden layers with a different number of neurons, the optimal number of layers, nodes and optimizer were determined by using a hyperparameter optimization, that is configured to minimize the MAE of resulting predictions and based on a TPE sampler. As a result the following architecture is chosen: One input layer, two hidden layers, and one output layer. The descriptive

TABLE 3. Hyperparameters for ANN predicting WP. Number of layers, corresponding nodes and optimizer are selected by optuna optimization.

Parameter	Selected Values
Input Layer	2
Output	length of forecast
Training algorithm function	Adam [29]
Hidden Layer 1 - Dense - function	relu
Hidden Layer 2 - Dense - function	relu
Hidden Layer 1 - Dense - size	16
Hidden Layer 2 - Dense - size	16
Weight update method	Batch- mode
Early Stopping	patience:2
Learning rate	1E-3
batchsize	150
Loss function	MAE

features are wind speed and temperature. The corresponding hyperparameters are shown in Table 3.

D. FORECAST OF N-1 SECURITY MAXIMUM LINE LOADING

Congestions occuring in the high voltage grid are determined considering the n-1 security [23], [24]. Thus, the ANN is trained with the maximum load on each line resulting from all possible contingencies. Transformer injections into the 110 kV distribution grid are used as input features, and as target feature the maximum loading of the components. The maximal loading for each line, occurring during contingency analysis, is determined via DIgSILENT PowerFactory. The data of vertical power flow on the MV/HV transformers is described in Section III-A2. A description of the used distribution grid can be found in Fig. 7 of our previous publication [14]. In the contingency analysis the load flow for every outage is calculated and in the end the maximal loading for each line is selected. As outages the lines of the 110 kV distribution grid and EHV/HV transformers are considered.

The ANN model consists of one input layer, a number of dense layers and an output layer. The number of inputs is equal to the amount of transformers, whereas the number of outputs is defined by number of lines. The other hyperparameters are determined by a hyperparameter optimization. The optimization model is configured to minimize the MAE of predicted loadings and is based on a TPE sampler. The values chosen for training are shown in Table 4. As a loss function the *mean absolute error* is chosen and *accuracy* is applied as metric.

E. ERROR MEASURES

$e_t = y_{\text{true}} - y_{\text{predicted}}$ describes the error of a prediction, where y_{true} denotes the actual value and $y_{\text{predicted}}$ the prediction. In order to validate the presented approaches the following error measures are used. For time series forecast measure Mean Absolute Error (MAE), Mean Squared Error (MSE), Relative Mean Absolute Error (RMAE), and Root Mean Square Error (RMSE) are applied. The MAE is defined as:

TABLE 4. Hyperparameters selected by optuna optimization for ANN predicting maximum line loading.

Parameter	Selected Values
Input Layer	Number of nodes
Output	Number of lines
Training algorithm function	Adam [29]
Hidden Layer 1 - Dense - function	relu
Hidden Layer 2 - Dense - function	relu
Hidden Layer 3 - Dense - function	relu
Hidden Layer 4 - Dense - function	relu
Hidden Layer 5 - Dense - function	relu
Hidden Layer 1 - Dense - size	88
Hidden Layer 2 - Dense - size	56
Hidden Layer 3 - Dense - size	24
Hidden Layer 4 - Dense - size	120
Hidden Layer 5 - Dense - size	88
Weight update method	batch-mode
Early Stopping	patience: 2
Learning rate	0.008
clipnorm	0.5
clipvalue	0.7
batchsize	150
Loss function	MAE

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{4}$$

where n denotes the number of values. MSE can be calculated as:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}e_i^2 \tag{5}$$

RMAE describes the MAE of the considered model in relation to MAE* of a reference model:

$$\text{RMAE} = \frac{\text{MAE}}{\text{MAE}^*} \tag{6}$$

The RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}e_i^2} \tag{7}$$

A detailed description of the error measures is given in [32].

For the classification task following error measures are applied, described in [33]. Precision which is calculated as:

$$\text{precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{8}$$

recall, that is described by:

$$\text{recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{9}$$

and F_1 score which can be calculated by:

$$F_1 = \frac{2\text{TP}}{2\text{TP}+\text{FN}+\text{FP}} \tag{10}$$

In the equations (8-10) TP denotes true positives, FP denotes false positives, and FN denotes false negative values.

TABLE 5. Python packages and the corresponding field of application.

Python package	Version	Application
TensorFlow [34]	2.8.0	ANN model design and training
Optuna [30]	2.10.0	ANN parameter optimization
pandas [35]	1.4.2	data processing
NumPy [36]	1.22.3	data processing
Scikit-learn [37]	1.0.2	error metrics, scaling
statsmodel [38]	0.13.2	ARIMA models
pmdarima [39]	1.8.5	ARIMA parameter optimization

III. RESULTS

The calculations were performed on 64-bit windows machine with an Intel©Core i5-6500 CPU @ 3.2GHz and 16 GB RAM. Power Flow simulations were calculated with DIgSILENT PowerFactory and the programming language was Python 3.9. The relevant Python packages used are shown in Table 5.

In order to validate and compare the proposed ANN models, we used an additional model for evaluation. Makridakis et al. proposed models for comparison such as, among others, the naïve model and ARIMA [40]. The parameters for fitting the ARIMA model were obtained by analyzing the autocorrelation and partial autocorrelation of our data and by an Augmented Dickey-Fuller test.

For the vertical power flow on MV/HV transformers the augmented Dickey-Fuller test (ADF) yields a test statistic of -17.09 and a p-value of 0.0 therefore, it can be assumed that the data are stationary. The Ljung-Box chi-square statistic results a p-value of 5.66, which is greater than the significance level of 0.05 and therefore indicating that the residuals are independent. After analysing the auto correlation function (ACF) and the partial auto correlation function (PACF) we used a parametrization of (1,1,0) for the ARIMA model. The resulting time series, predicted for the next 12- time steps showed a poor accuracy. Also, varying the parametrization did not improve the results. The same process was performed to fit an ARIMA model to predict the maximum line load considering n-1 security. Nonetheless, these predictions also showed poor accuracy. However, since the fitted ARIMA model gave poor results and performing further parameter optimizations did not yield promising results either, we decided to use the naive approach as a comparative model.

A. DATA

For simulation, the following data were used for a region in the north of Germany in 2015 and 2016:

- Temperature and wind speed data for a height of 73 m with a hourly resolution are taken from the COSMO-DE analysis of the DWD (Deutscher Wetterdienst) [41]. Our simulation time step is 15 minutes, so the data has been interpolated. In our previous work [13], it has been shown, that the error is acceptable.
- Global irradiance, based on Meteosat-SEVIRI data [42] with a temporal resolution of 15 minutes.
- Historic transformer power of the MV/HV transformers with a temporal resolution of 15 minutes provided by the corresponding DSO.
- Transformer specific historic curtailment given for the exact minute provided by the corresponding DSO.
- Calculated wind power data per transformer, which contains the generated power of all wind parks connected to a HV transformer with a time resolution of 15 minutes. The calculation is performed with an adjusted wind turbine power curve and the wind speed data from DWD for the considered years. A further description can be found in Section II-B.1 of our previous paper [14].
- Maximum loading on each component was determined by load flow calculations taking into account contingencies in DIgSILENT PowerFactory. The time series used has a time step of 15 minutes and was considered for the period March-December 2016. As input for each node, corresponding to the node type, the preprocessed transformer data (see Section III-B) and the wind power data (mentioned above) were used. The resulting line loadings are shown in Fig. 12.

1) FEATURE SELECTION

The features initially considered were selected based on our expertise gained from our previous work [13], [14], which pursues the same objective using different methods. Separating the electricity grid into its components, the required features can be listed: As the vertical power flow on MV/HV transformers is determined by its consumers and generators, where RE account for the largest share, time is selected in order to represent load and appropriate meteorological data is chosen, like wind speed, temperature and global irradiance, to represent the generation of PV and wind power. As target feature the measurements of the transformers is applied. The selected features for the component loading prediction considering n-1 security, are based on data, which is required for a contingency analysis performed in PowerFactory: as input the power-injections of all nodes and as target feature the resulting maximum loading of all components in the considered grid is taken. For the final selection of features the correlation coefficients between descriptive features and target features have been considered. The selected features are listed per ANN model in Table 6.

2) PREPROCESSING TRANSFORMER POWER MEASUREMENTS

As the power data of MV/HV transformers were provided as measurements, curtailed power values are included. Thus, at times of curtailment, low power values can be found at high wind speeds (>15 m/s) instead of expected high power values. Fig. 4 visualizes this effect exemplary for one transformer. It shows transformer power over wind speed at times of curtailment. The upper subplot represents raw data, where negative power values (which represent demand/power

TABLE 6. Visualization of the selected features per ANN model. Descriptive features are assigned as 'input' and target features as 'output'.

Features	ANN Models		
	Vertical power	Wind power prediction	n-1 loading
Wind speed	input	input	
Global irradiance	input		
Temperature		input	
Transformer power	input/output		input
Wind power		output	
Max loading			output
Sin(day of the year)	input		
Hour	input		
Cos(time)	input		

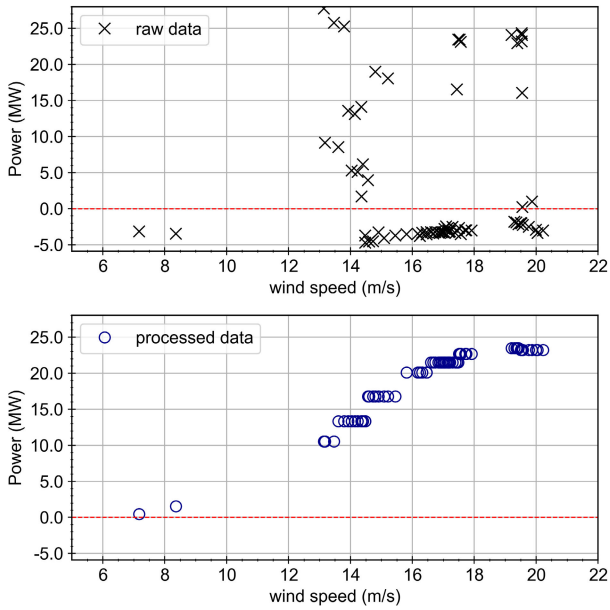

FIGURE 4. Visualization of transformer power over wind speed shown for an exemplary transformer at times of curtailment. The right subplot shows raw data, and the left processed data.

injections into the MV grid level) can be found at high wind speeds. In order to identify congestions, the potential transformer power should be represented in the data. Therefore, curtailed power values have to be replaced by those of potential power without curtailment. The lower subplot of this figure represents the processed data. Here the power increases with increasing wind speed as it is expected.

The processing is done by a look-up table, assigning available transformer power to each wind speed, excluding power at times of curtailment. The values of one exemplary look-up table are shown in Fig. 5.

3) PREPROCESSING (TIME)

As consumers are connected to the transformers, time and date play a certain role for the load profile. In order to use timestamp as a descriptive feature, it has to be transformed.

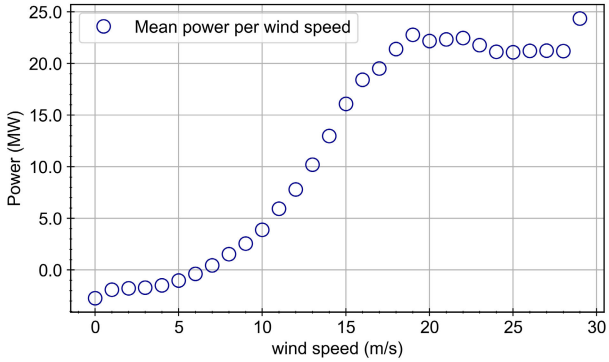

FIGURE 5. Visualization of the average transformer power binned per wind speed where values at times of curtailment have beeen excluded.

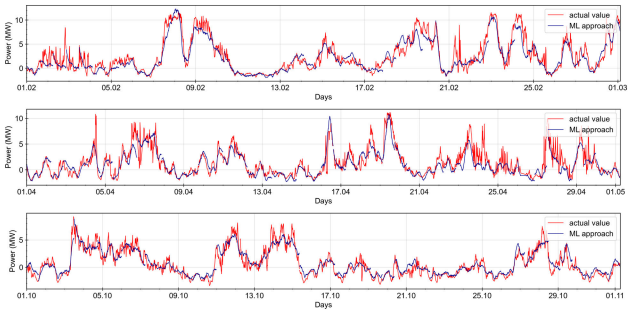

FIGURE 6. 15 hour forecast of MV/HV transformer power (of the cluster *mainly backfeed*) calculated for a month, shown for three sets of a five-fold cross-validation.

The timestamp is given as an information of weekend and the date (with the information of season) as sinus of the year. The time is represented by a corresponding sinus and cosinus value.

B. VERTICAL POWER PREDICTION

The neural network predicting the vertical power is applied for an exemplary transformer of the transformer cluster *mainly backfeed*. In Fig. 6 the predictions and actual values over time for one month are shown for three sets of a five-fold cross-validation. These forecast horizon is set to 15 hours which is an equivalent of 60 timesteps. Comparing the ML prediction and target values it is obvious that the measurements can be predicted with some deviations. The predictions for the five datasets have an average MAE of 0.93 (the minimum MAE of all five data sets is 0.78 and the maximum is 1.05) which is 3.9% of the maximum power measurement of the corresponding transformer for years 2015-2016.

In order to evaluate the effectiveness of the presented approach, the ANN for vertical power prediction is compared to a Naive-Forecast approach, using the previous values to predict the future ones. So to predict the next 60 steps ahead, the values of the last 60 time steps are applied. The corresponding error measures are shown for each vertical power direction (backfeed and injection) in Table 7. Following error

TABLE 7. MAE comparison between the ANN- and the Naive-approach. Calculated each for backfeeded power (into the HV-grid level) and injected power (into the MV-grid level).

power direction	ANN		Naive		Comparison
	MAE	MSE	MAE	MSE	RMAE
both	0.93	2.06	2.49	12.2	0.37
backfeed	1.16	2.84	2.95	15.05	0.39
injection	0.59	0.99	1.76	7.7	0.34

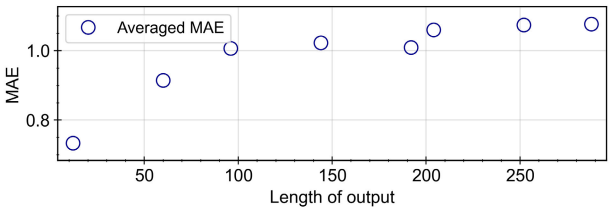

FIGURE 7. Averaged MAE (gained by multiple training) per forecast horizon.

measures are considered: MAE, MSE, and RMAE for which the Naive model is used as reference.

The RMAE of 0.37, describing the relation between the ANN and the Naive approach, demonstrates the superiority of the ANN in relation to accuracy. For further validation a unpaired two-sample-t-test was performed. As the resulting p-value was smaller than 0.01 we reject the null hypothesis and conclude that the means of both error values are statistically different.

In Fig. 7 the impact of the predicted output length on the accuracy is visualized. To this end ANNs have been trained multiple times for each forecast horizon, which was set in discrete steps between 12 and 288 steps ahead. As expected the MAE shows the tendency to increase with increasing output. The error growth shows an asymptotic behavior. One exception can be seen at the mean MAE for ANNs trained with a forecast horizon of 192 time steps, which show a slightly lower MAE than these with a forecast horizon of 144.

In the following, ANNs are trained for all transformers of the model region with an output length of 12 time steps. Again cross-validation is performed in order to ensure the stability of the trained model. To get a better overview of the overall performance the analysis is performed on the basis of clustered transformers.

The resulting MAEs in MW for the calculated vertical transformer power are shown cluster vise in Table 8. The MAE over all transformers ranges from 0.4 MW MAE to a MAE of 2.72 MW for injected power and from 1.0 MW MAE to a MAE of 3.85 MW for backfeeded power. This range of accuracy can be attributed to the different vertical power flow characteristics of the transformers (see section II-B). It can be seen that the transformer cluster *Exception* has the smallest averaged MAE (2.06 MW) for power backfeed into the HV grid level. Whereas, for power injection into the MV grid level the transformer cluster *demand* has the smallest average MAE (0.41 MW). Comparing the MAEs

TABLE 8. The transformer MAEs calculated with a time horizon of 12 steps for the ANN prediction of vertical power shown separately for injections into the MV grid level and backfeed into the HV grid level. The maximum, minimum and average MAE of the averaged cross-validation test results is given for each transformer cluster.

Cluster	MAE [MW] power backfeed			MAE [MW] power injection		
	max	min	av.	max	min	av.
Exception	2.91	1.41	2.06	nan	nan	nan
backfeed	3.85	1.21	2.15	2.24	0.56	1.23
mainly backfeed	3.37	1.00	2.19	1.85	0.49	1.16
mainly demand	3.26	1.57	2.40	2.72	1.19	1.83
demand	2.31	2.01	2.16	0.42	0.40	0.41

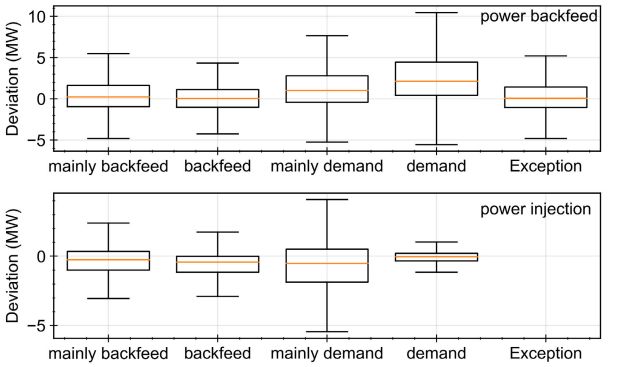

FIGURE 8. Comparison of deviations shown per transformer cluster. The subplot with power backfeed represents deviations at times of actual power flow from MV to HV grid level. Accordingly, the subplot with power injection represents deviations at times of actual power flow from HV to MV grid level.

for the predicted backfeeded power with the injected one, it can be seen, that the error measures are smaller for power injections (maximum average MAE is 1.83 MW) than they are for the backfeeded power (maximum average MAE is 2.4 MW). It has to be considered that the absolute maximum power deviates between the transformers. Further, the power injected into the MV grid tends to be smaller in absolute values, than the backfeeded power.

An overview of the deviations between predicted and actual transformer power is given in Fig. 8 where deviations are shown per transformer cluster. It can be seen, that the power injections into the MV grid level are predicted with lower deviations than those of power backfeed. The differences between the clusters become obvious in the distributions of the deviations, especially for the deviations of injected power: Where the distributions of the clusters *mainly backfeed* and *backfeed* are quite similar (in addition, the *backfeed* cluster tends to have a lower spread of deviations than does *mainly backfeed*), but the distributions of *mainly demand* show a wider range and those of *demand* a much smaller one than all of them.

Over all clusters the errors are smaller for predicted power injections. It can be seen, that those clusters that have clear tendency for backfeed or injection (*backfeed*, *demand*) show the smallest error and deviations. For both clusters, there are

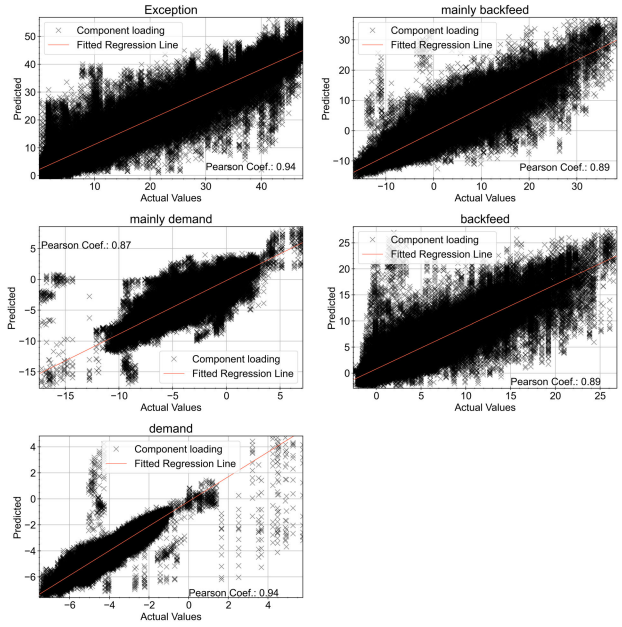

FIGURE 9. Regression plots shown per exemplary transformer of each cluster.

more cases of underestimation of actual power (shown as negative deviation), while for the clusters *mainly demand*, *mainly backfeed* predicted power is more often overestimated (shown as positive deviation).

Furthermore, the distributions of deviations indicate a tendency of errors being smaller for backfeeded power at transformers that have smaller demand than generation. A backfeed of power takes place when a surplus of RE power is presented. The main share of RE power in the model region comes from WP, which has a logistic relationship with wind speed. Therefore, it is assumed that backfeeded power, where WP generation is higher than demand, can be predicted more precise than injected power, due to this simple relationship between wind speed and generated power. Accordingly the clusters *demand* and *mainly demand* show greater inaccuracies at times of backfeeding power.

Another option to analyse existing deviations between predicted and actual values is to look at corresponding regression plots. The correlations shown in Fig. 9, deviate between 0.87 and 0.94, which indicates a good relation between predicted and actual values. The plotted data are scattered over a wider range. Bigger deviations from the diagonal represent values where the neural network has not learned the relation between input and target values properly. Deviations that occur at actual values around 0 (as seen in the example of the *mainly backfeed* transformer and also the *backfeed* transformer) indicate that the input parameter is associated with a higher target value at these times than the backfeed actually was. Reasons for that behavior could be either a curtailment of RE generation due to e.g. bat protection or negative price, or the load has been unexpectly high. Generally wider deviations indicate that the model has not learned the

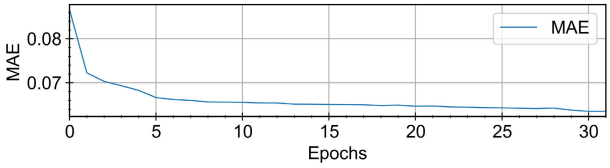

FIGURE 10. Learning curve of an exemplary transformer of the cluster ***Exception*** and a forecast horizon of 12 time steps.

relation between input and target values for all cases, so for future work further adjustment is required.

A learning curve of an exemplary transformer of the cluster *Exception* is shown in Fig. 10. It visualizes a sufficient performance of the learning process.

During the review process of this article, we came across another paper that also proposed a LSTM for predicting vertical power on MV/HV transformers [9]. Both models were developed independently of one another. The approach proposed in [9] uses a similar architecture for the LSTM, with two separate input layers, one for vertical power flow measurements and the other for meteorological data and information about the day. There are the following differences between the two approaches: Instead of connecting both input layers with LSTM layers, the input layer for the meteorological features of our proposed model is connected with a dense layer. Further, in our approach, the temporal features are connected to the same input layer as the vertical power measurements and not to the input layer that receives meteorological features. And the LSTM proposed in [9] uses more meteorological information (see Section III-B in [9]) than our method: Wind speeds for 10 m and 100 m altitude; temperature and dew point temperature; predicted albedo and surface downward solar radiation, surface pressure, and total precipitation.

Another difference is the number of nodes per layer: the maximum number of nodes per layer for our LSTM is 128, while the other model has 100 per LSTM layer and 500 per dense layer.

In the following we compare one proposed approach in [9] (‘‘LSTM_updated’’) with our approach in terms of the published normalized RMSE and the Pearson correlation coefficient. Accordingly, we normalized the RMSE as Brauns et al. The normalization is described in equation 2 in [9]. The RMSE was calculated for each transformer separately. The forecast horizons considered in [9] are 3-, and 15 hours. The whiskers of the resulting boxplot over all normalized RMSE values is given in Table 9. It can be seen that the normalized RMSE of our proposed model is slightly lower than those of the comparison model. The ANNs show values of 0.16 for the top whisker at a forecast horizon of 15 hours. Whereas, the ‘‘LSTM_updated’’ of [9] shows values between $]0.29 - 0.31[$ for the top whisker for the forecast horizon of 4 hours. As already shown in Fig. 7, an increase in the errors can be expected with increasing forecast horizon. So the results indicate, that our approach gains slightly higher accuracy in the case of normalized RMSE.

TABLE 9. Rough comparison of normalized RMSEs for ANN models predicting vertical power on MV/HV transformers for different forecast horizons. Since the RMSE error measures were given only as boxplots in Fig. 7 in [9], ranges are given instead of concrete values for the whiskers.

Model	forecast horizon	whisker bottom	whisker top
ANN - Vertical Power	15h	0.08	0.16
ANN - Vertical Power	3h	0.05	0.13
[9] - LSTM_updated	16h	~0.1	0.34>x>0.3
[9] - LSTM_updated	4h	0.1>x>0.09	0.31>x>0.29
[9] - LSTM_updated	1h	0.095>x>0.08	~0.25

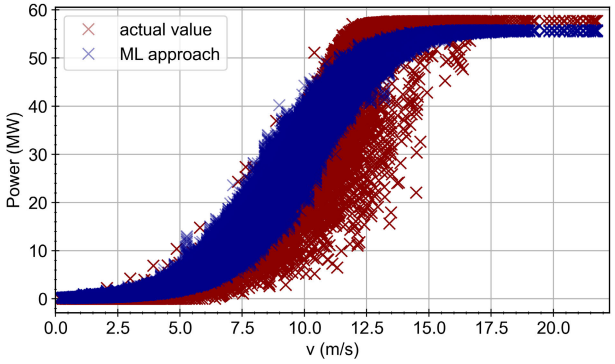

FIGURE 11. Predicted wind power via ANN over corresponding wind speed compared to actual values. The prediction horizon was 12 time steps (3 hours).

In [9] the Pearson correlation coefficient is given for 'transformer 5' for 4 different forecast horizons: 1 h, 4 h, 16 h, 48 h (see Fig. 6 in [9]). The LSTM with regular training shows a correlation of 0.9 for a 4 h and 0.91 for a 1h forecast horizon. Comparing those correlation coefficient to the ones, shown in Fig. 9, with a forecast horizon of 3 h it shows a similar accuracy as our proposed model, which ranges between [0.87-0.94].

C. GENERATED WIND POWER PREDICTION

In order to validate the ANN trained to predict the generated power of wind farms connected via transformers to the HV grid level, the predicted WP is plotted over its corresponding wind speed, see Fig. 11. Here the corresponding ANN was trained to predict the next 12 time steps for an exemplary transformer. The relationship between wind speed and WP can be described by a logistic curve. Since the scatter plot follows this pattern, it can be concluded that the trained ANN has learned the relationship between wind speed and generated power.

For further validation error measures of the predicted power are compared to the outcome of the Naive forecast approach. In Table 10 the resulting MAEs, MSEs, and RMAE, for which the Naive approach is used as reference model, are contrasted. Comparing the MAEs it can be seen that the ANN outperforms the Naive approach wherefore, the RMAE confirms the superiority of the ANN with a value

TABLE 10. Comparison of error measures between ANN-predicting WP of a wind park connected to a HV-transformer and the Naive-approach.

ANN		Naive		Comparison
MAE	MSE	MAE	MSE	RMAE
3.25	26.96	6.09	83.45	0.53

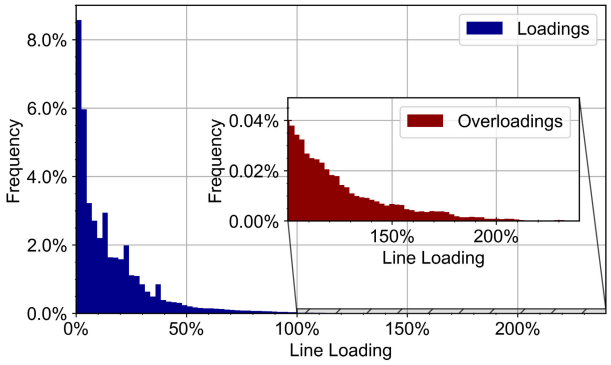

FIGURE 12. Histograms of the line loadings resulting from load flow calculation considering contingencies. The main graphic shows the whole data set and the overloadings are visualized in the zoomed area.

of 0.53. Again an unpaired two-sample-t-test was performed. With a p-value smaller 0.01 the null hypothesis was rejected. Thus, we conclude that the mean error values of the ANN and the Naive model are statistically different.

D. PREDICTION OF LINE LOADING CONSIDERING N-1 SECURITY

An ANN is trained to predict the maximal loading of the components of a realistic 110 kV distribution grid located in the north of Germany (The topology is shown in Fig. 1. For a more detailed description see Fig. 7 and subsection III-A in [14]) for contingency cases. Data from March of 2016 until December 2016 in a 15 minute time resolution was used. The distribution of the resulting line loadings are shown in Fig. 12. The loadings have a mean of 16.54% and a standard deviation of 20.76%. It can be seen that most of the time (in 93.77% of all cases) the loading is smaller than 50%. In total only 1.07% of the data show overloadings (loadings with a value equal or greater than 100%), which can reach a maximum value of 239.79%.

In order to describe the gained precision the predictions are analyzed regarding the predicted line loading accuracy on the one hand and on the other hand the capability to detect congestions leading to curtailment. In Fig. 13 the predicted n-1 security loading is compared to the actual loading for two 110 kV lines and one HV/EHV-transformer. The prediction was made for a forecast horizon of 12 time steps (3 hours). Regarding the course of the actual loading, the ANN shows a high match, with only minor deviations.

In Table 11 error measures are compared against a Naive approach considering a forecast horizon of 12 steps. Error measures are calculated for each component. In order to give

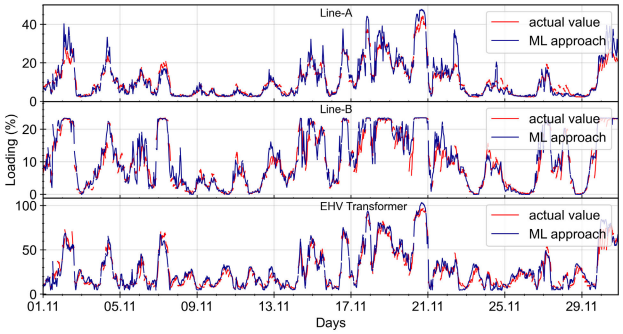

FIGURE 13. Loading for two 110 kV lines and one HV/EHV transformer over time.

TABLE 11. Comparison of error measures between ANN-predicting n-1 security component loading and the Naive-approach. The mean, maximum and minimum error values of all components are shown.

	ANN		Naive		Comparison
	MAE	MSE	MAE	MSE	RMAE
mean	2.59	22.82	4.89	97.87	0.53
max	8.18	143.15	17.45	739.06	0.47
min	«0.01	«0.01	«0.01	«0.01	-

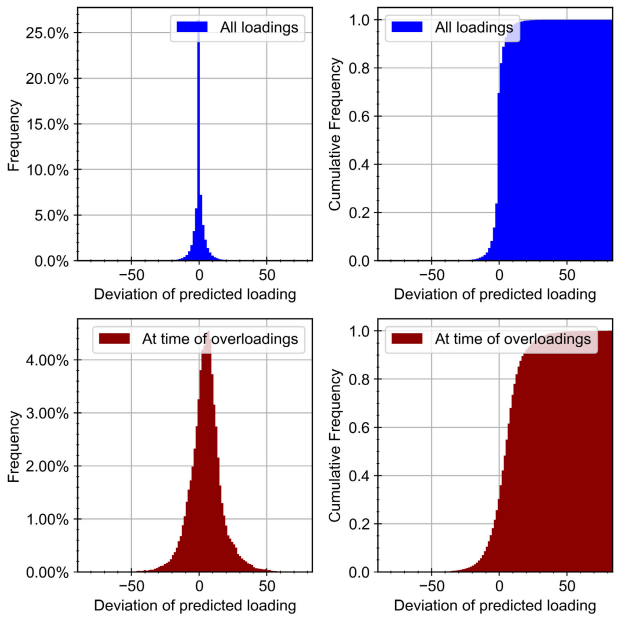

FIGURE 14. Histogram and empirical cumulative distribution of the deviation between the actual component loading and the predicted one. The upper subplots shows the distribution of the whole dataset, while the lower subplot shows them for times of actual congestions.

an overview mean, maximum and minimum of the error measures are given. The values of RMAE demonstrate that the ANN shows the better performance than the Naive approach. The ANN has a MAE of 2.59, which is quite small considering the fact, that the predicted loading represents already the n-1 security case. Again an unpaired two-sample-t-test was performed for each grid component individually. The null hypothesis was rejected ($p < 0.01$) for 87.5% of the grid components. No statistically significant difference in MAE could be shown for the remaining 12.5% of grid components.

Taking only loading at times of congestions into account (loading higher 100%), the corresponding average MAE is 8.77, which indicates higher deviations for the predicted loading at times of congestion. Regarding curtailment the effect of a higher error depends on whether the predicted loading is over- or underestimated. In event of an overestimation, the effects would not be so severe, since it has to be curtailed anyway and thus grid security would be maintained. In the worst case only the amount of curtailment could be end up higher than actually needed. Whereas in case of an underestimation, required curtailment would not be detected. In order to determine the occurrence of under- and overestimation the distribution of the deviations are analyzed in the following.

In Fig. 14 the distribution of deviation between the predicted and the actual component loading is shown. The distribution is nearly symmetrical with a mean of 0.09 and a standard deviation of 4.77, showing a little tendency to overestimate the components loading (represented by negative values). 84% of the predicted loadings show deviations smaller than 5% loading. So overall it can be said that the model is suitable for predictions of loadings considering n-1 security.

Regarding the capability of predicting congestions (components loaded higher than 100%) following statements can be made: Deviations of predicted loadings for actual values higher 100%, appear with a mean of 5.08 and a standard deviation of 11.96. The corresponding distribution shows that the predicted loadings have a slight tendency to underestimate (represented by positive values) at times of actual congestions. This means that the event of undetected congestions can occur more often. (33% of the values have an error smaller 5% and around 10% of the values an error bigger 20%).

Focusing on the accuracy regarding the determination of curtailment requirement, three metrics are evaluated: precision, recall, and F_1 score. As the ANN predicts the resulting component loading considering contingencies, it can also be applied to determine curtailment requirement. Defining a component with a loading bigger than 100% as a congested component the accuracy for the congestion forecast can be evaluated. The results are shown in Table 12 and compared for validation to the prediction model proposed in [12]. In this study Staudt et al. proposed an ANN trained for the hourly prediction of congestions occuring in transmission grids. It can be seen that for all metrics, the presented approach of this study shows higher values than those presented in [12]. But the difference between the predictions (transmission grid and hourly predictions in [12] vs. distribution grid and quarterly predictions in the approach of this study) can also have an impact on the achieved accuracy. At least it can be stated that the approach is suitable to predict congestions.

The metrics were also calculated separately for each voltage level of the operating equipment. Comparing the metrics

TABLE 12. Accuracy metrics for the determined congestion according to the n-1 security line loading prediction (with a prediction horizon of 12 times steps) compared to the congestion prediction proposed in [12]. Further the metrics are compared for the TSO and the DSO components.

model approach	precision	recall	F_1 score
[12]	0.70	0.64	0.65
n-1 security ANN	0.92	0.92	0.92
n-1 security ANN - *TSO*	0.91	0.77	0.84
n-1 security ANN - *DSO*	0.92	0.92	0.92

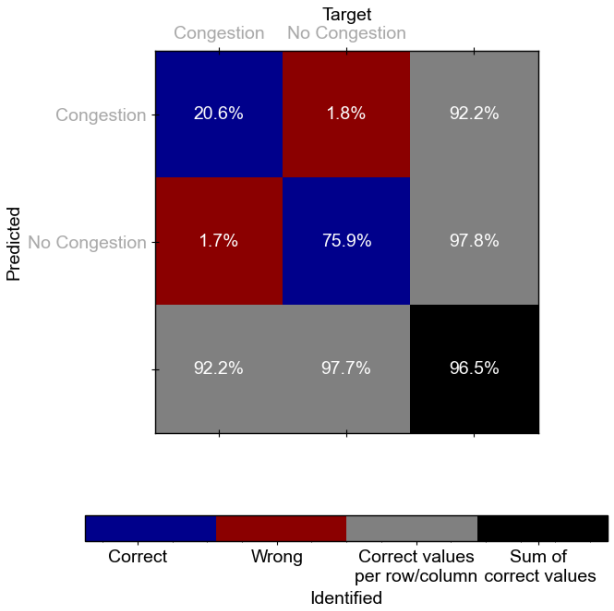

FIGURE 15. Confusion matrix.

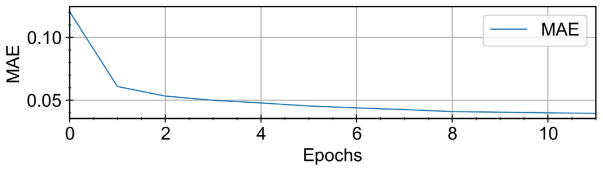

FIGURE 16. Learning curve of the ANN predicting line loading.

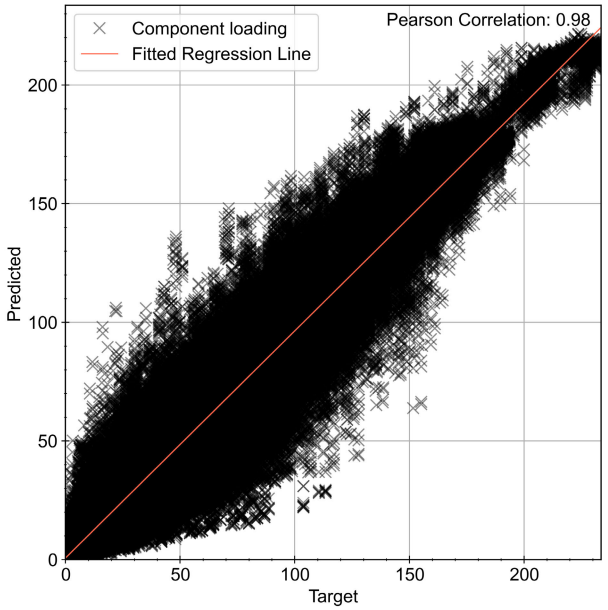

FIGURE 17. Regression shown for target and predicted loading values.

for the components of the respective grid operator, it can be seen that the model achieves higher accuracy for the DSO components. This could be caused by the fact, that there are more DSO-components (which make 93% of all components), than TSO-components (which make 7% of all components). Therefore, the ANN has more opportunities to adjust the weights in a way, that they predict the DSO-component loadings with smaller error, than they have for the TSO-component loadings. It can be concluded that higher errors have to be expected for predicted TSO congestions, than for predicted DSO congestions.

In Fig. 15 the confusion matrix of the predicted congestion is shown for the trained ANN line loading forecast model. It shows that 96.5% of the values were correctly determined. Congestions were correctly predicted in 92.2% of actual cases. At times without congestion, a normal grid condition was correctly forecasted for 97.8% of the cases. These results indicate that this ANN is a promising approach to detect congested grids considering n-1 security. The learning curve for this ANN is shown in Fig. 16. It shows that the learning process has performed sufficently.

The regression plot, shown in Fig. 17, has a correlation coefficent of 0.98, indicating a good correlation between target and predicted values. The wider deviations between predicted and target values occur mainly between an actual loading of 0% and around 150%. At actual line loading greater 190% the deviations are smaller. This indicates that the high loading events can be predicted with smaller errors than for lower loading events.

In order to analyze the model performance, cross-validation was also performed by splitting the data into 5 sets and maintaining the temporal order. The resulting error measures are presented in Table 13. As the averaged F1-score is 0.78 instead of 0.92 emphasizes the impact of the training data on the resulting accuracy. One reason for this effect could be the fact, that training or test data have too few congestions due to data splitting.

Table 14 shows the distribution of F_1 scores determined for each component individually. The impact of a generation unit on a congested line is taken under consideration for the selection of which unit needs to be curtailed. Hence, the potential to predict bottlenecks for the correct line is crucial. A high accuracy regarding location of predicted congestions is given at high values of component specific F_1 score. For the components an average F_1 score of 0.8 is reached, which

TABLE 13. Averaged accuracy metrics for the determined congestion according to the n-1 security line loading prediction (with a prediction horizon of 12 times steps) considering cross-validation compared to the congestion prediction proposed in [12]. Further the metrics are compared for the TSO and the DSO components.

model approach	precision	recall	F_1 score
[12]	0.70	0.64	0.65
n-1 security ANN	0.85	0.75	0.78
n-1 security ANN - TSO	0.99	0.19	0.28
n-1 security ANN - DSO	0.85	0.75	0.78

TABLE 14. Statistic of individual F_1 scores for each congested grid component.

	Descriptive		Extrema		Quantiles		
parameter	mean	std	max	min	25%	50%	75%
value	0.8	0.12	0.94	0.52	0.73	0.82	0.91

TABLE 15. The RMAEs shown per transformer cluster considering cross-validation, with the model presented in [13] as reference.

Cluster	RMAE power backfeed			RMAE power injection		
	max	min	av.	max	min	av.
Exception	0.82	6.13	0.89	—	—	—
backfeed	0.69	0.62	0.57	0.16	1.87	0.29
mainly backfeed	0.46	0.64	0.64	0.52	0.64	0.66
mainly demand	0.24	0.36	0.32	0.31	0.70	0.55
demand	0.39	2.79	0.57	0.1	0.65	0.21

indicates that the model detects congestions for the correct components with sufficient accuracy.

E. COMPARISON WITH PARAMETRIC APPROACH

In our previous works we presented parametric approaches for vertical power flow predictions on MV/HV transformers [13] and a curtailment forecast for distribution grids [14]. The vertical power flow in [13] is calculated by aggregating generated and consumed power for each MV/HV transformer. The generated PV power is calculated via characteristic power curves and global irradiance for transformers coordinates. The corresponding WP power is defined by an adjusted function as it is the load profile. Both are adjusted by using the identified generated WP power and consumption from measured transformer data. By using the according time and meteorological data as input values to these adjusted functions, the corresponding MV/HV transformer power is calculated.

The curtailment prediction in [14] is realized by using the vertical power flow (calculated according to [13]) as backfeed into a 110 kV distribution grid. Then with power flow calculations considering contingencies congestion are detected and required curtailment is determined. Both approaches have been validated with the same data as the ANNs presented in this study.

In order to validate whether the ANNs or the parametric methods are superior, the methods are compared in the following: First the RMAEs are calculated for each transformer cluster, with taking the model from [13] as reference model. As in Table 2 of [13] different combinations of methods are analyzed, these MAEs are taken for comparison, which have the lowest value.

TABLE 16. Accuracy metrics for the determined congestion according to the n-1 security line loading prediction compared to the congestion prediction proposed in [14].

grid level	F_1 score [14]	F_1 score ANN
Aggregated	0.84	0.92
TSO	0.79	0.84
DSO	0.47	0.92

The ANN show lower values of MAE. Outliers are the minimum MAEs for the clusters *Exception* and *demand* of power backfeed and the minimum MAE for the cluster *backfeed* of power injection. This indicates that the proposed ANNs achieve higher accuracy and are regarding error measures superior. Comparing the transformer clusters, the cluster *Exception* show higher RMAEs for backfeeded power than the others. With minimum value greater 1 and a RMAEs near to 1 (0.89) describing the relation of the average MAE, it indicates that the proposed ANNs show for some cases only slightly better accuracy and for some transformers of the cluster *Exception* even worse. For these transformers more research is required to adjust the corresponding ANNs.

In order to evaluate the performance of congestion detection the F_1 score is compared in Table 16 between the prediction of n-1 security line loading presented in this study and the *6m* scenario presented in [14] (see Table 2 of [14]). The comparison shows that the presented ANN outperforms the method in [14]. Looking at the F_1 scores for the individual voltage levels shows differences between both methods in terms of accuracy of congestion prediction. While the method presented in [14] tends to predict congestions on TSO-components with a higher precision, the proposed ANN achieves higher accuracy for DSO-components. One reason could be that the frequency of congestion per voltage level is decisive for ANNs as mentioned earlier. Whereas, for power flow calculations the arriving aggregated power at the HV/EHV transformers increases the probability to detect congestions correctly. However, one restriction regarding the evaluation of the results has to be taken into account: While for the ANN true congestions are defined by the results of load flow calculations (at times of line loading greater 100%), for the method in [14] true congestion are defined by actual curtailment. Some deviations in the predicted congestions could be due to the fact that the behavior of grid operator is not considered in parametric curtailment determination (as discsussed in [14]).

F. RUN TIME ANALYSIS

For further information, the models proposed in this paper were analyzed in terms of their computation time for training and prediction. Measurements are taken for the ANN models for vertical power prediction, WP prediction, and the ANN

TABLE 17. Computational time for training and prediction of ANN models and the contingency analysis in PowerFactory given in seconds. Time measurements are taken 20 times.

Model	step	min [sec]	mean [sec]	max [sec]	std [sec]
ANN - Vertical Power	train	13.03	18.05	23.37	3.27
ANN - Vertical Power	predict	1.42	1.45	1.53	0.03
ANN - WP	train	0.81	0.92	1.17	0.1
ANN - WP	predict	0.11	0.12	0.16	0.01
ANN - component loading	train	10.58	18.28	33.15	5.98
ANN - component loading	predict	0.39	0.54	0.7	0.1
Power Flow calculation	n-1	9.19	10.05	24.77	3.37

predicting component loading considering the n-1 security. In addition, measurements are also taken for contingency analysis (indicated in the Table as n-1), which is performed with DIgSILENT PowerFactory. The time measurements, which has been performed on CPU, is taken 20 times each and shown in Table 17.

As expected, the training time for the ANN-WP is by far the shortest. These comparatively short times can be explained by the relatively simple ANN structure (fewer nodes and layers than the other two models). In addition, the relation between input and target features that need to be learned is simpler. The ANN for vertical power prediction and the one for component loading forecast have almost the same training time on average, but the standard deviation of the second model is much higher.

Comparing the times needed for calculating line loading considering n-1 security, it can be seen that the ANN shows a much smaller computational time in contrast to the contingency analysis performed via DIgSILENT PowerFactory. Training of the ANN takes on average with 18.28 seconds about 0.55 times as long as the contingency analysis.

IV. DISCUSSION

The presented ANNs for vertical power prediction on MV/HV transformers and the forecast of component loading considering contingencies contributes to the current need for congestion predictions, which is caused by an unbalance existing between RE integration and grid expansion, and recent changes in regulation regarding grid congestions (Redispatch 2.0).

The detection of overloaded components is only possible if potential infeeds into the distribution grid are known. Such injections are determined by the presented ANN approach for vertical power predictions. The approach achieves predictions with small deviations. It was determined that larger errors are to be expected with increasing forecast horizon. Thus, larger deviations have to be considered with regard to the day-ahead forecasts partly required for ''Redispatch 2.0''. Therefore, further research to establish a method to quantify the implemented uncertainty would be helpful. Furthermore, differences in the gained accuracy between different MV/HV transformer clusters have been identified. Especially these transformer clusters, which have more power backfeed into the 110 kV distribution grid (*Exception, backfeed, mainly backfeed*), tend to show lower errors for prediction of backfeeded power, compared to those which have more power injected into the 20 kV grid. More congestions are expected to be caused by backfeeds coming from transformers of these clusters. Thus, the errors for these clusters tend to be smaller. Therefore, the effect on the predicted congestion accuracy is expected to be reduced as well. Moreover, the difference in achieved accuracy between transformer groups highlights that the performance of ANNs is sensitive to the characteristics of the measured transformer power. Hence, it is recommended to further research regarding transformer cluster specific adjustments of the ANNs architecture.

As mentioned in the results, a LSTM predicting vertical power flow was published in [9] during our revision process. Those model has a similar ANN architecture to our model and their results show similar prediction accuracy to our approach. As both works were developed independently, this indicates that an architecture with two separate input layers is suitable for predicting vertical power flow. Brauns et al. showed that using an update strategy improves the performance of the LSTM [9]. This indicates that implementing an update strategy in our approach will also further improve the accuracy. Moreover, the model proposed in [9] uses more input features than our approach. As the results of both models show similar accuracy, our approach indicates that it is possible to achieve nearly equal accuracy with fewer features. Given that data is often difficult to obtain, our approach has the advantage that sufficient accuracy can be reached even with less features.

The transformers connected to the 110 kV distribution grid not only link MV grids, but also connect larger wind farms. As a high amount of power is feed into the grid by these wind farms, an ANN is designed and validated, which represents the generation of a wind farm. Solely relying on wind speed, temperature and the corresponding generated power, determined via a standard WTPC and scaled to the installed capacity, the approach can easily be utilized by other grid operators, that have no access to the information how much these wind farms generate but also need this information in order to determine the amount of curtailment required for their own power plants.

In terms of grid security it is crucial for grid operators to know whether there are upcoming congestions. Congestions are determined by considering contingencies. Accordingly one presented ANN predicts loadings resulting from n-1 security grid situations. Thus, occurring deviations in predictions do not have such severe consequences, since they already include critical cases. If the forecasts are used to predict curtailments, the slight tendency to overestimate loadings of the forecast could result in more redispatch being requested than necessary. But regarding the precision of predicting congestion, the analysis showed good efficiency detecting congestions. And results further indicate that the accuracy

of congestion prediction is higher for DSO- than for TSO components. Further, it was shown that the ANN has a high accuracy in predicting congestions for the correct component. Hence, the model could be used to select the required RE plants for curtailment needed to eliminate congestion.

Different methods to predict curtailment have been proposed in literature. In this study the presented ANN approaches for vertical power and line loading prediction are compared to corresponding parametric methods presented in [14] and [13], which applied the same data. The comparison shows that the ANNs provide more accurate results than the parametric methods with regard to the respective evaluation measures used. In case of the vertical power prediction one reason for the superiority of the ANNs could be that this method works directly with the data (transformer power), while the parametric method in [13] needs to dis-aggregate the transformer power in order to fit the functions describing WP or demand. As a result errors are included in the data used for fitting. These errors are avoided by using the transformer power directly in ANNs. Analysis showed that the ANN requires significantly less computation time for prediction of component loading considering n-1 security than load flow calculations. This makes the ANN also the more promising approach.

Following limitations of the study regarding field of analysis and characteristic of applied data have to be considered: The ANNs for vertical power prediction of MV/HV transformers were trained for transformers at which wind energy is the dominant renewable generation technology and has a high share of installed power. As the results already indicate, the different characteristics of vertical power flow lead to different accuracies. Further, the high share of installed WPs triggered several curtailment events. An unbalanced data set with a lack of special events like curtailments makes it more challenging for ANNs to learn the cause of such events and requires additional measures [43]. Furthermore, it must be noted that the ANN for predicting the component loading of an HV grid considering n-1 security was trained with data from a 110 kV grid connected to the EHV grid only by three HV/EHV transformers. Therefore, the influence of power injections coming from the EHV grid is limited. Another limitation is that only the maximum value of each n-1 security case is considered for a time step, rather than considering each contingency individually. The resulting simplification must be taken into account when evaluating the results.

Further research should concentrate on a method to quantify uncertainty of the predictions. Also ANN's architecture adjusted to each specific transformer should be investigated. A necessary step towards real time application is the analysis regarding model update strategies, especially required when the grid undergoes changes. These changes could be caused, for example, by an increase of installed RE power, a modification of consumption patterns, or the expansion of grid capacity. In these cases, the training data has to be regenerated and the ANN model has to be re-trained. If the grid topology changes in term of number of nodes or lines, the number of inputs, or the number of output neurons has to be adapted in the ANN for component loading prediction. The same adjustment has to be made, if the model is to be transferred to another grid region: The number of input and output neurons must be modified, and the ANN model must be trained with data representing the appropriate grid. Aiming at the improvement of predicting congestions in distribution grids, the next step would be to combine the ANN approach for vertical power prediction with load flow approaches to identify grid congestion and then to evaluate this hybrid method and its possible beneficial effects.

V. CONCLUSION

Increasing curtailment of RE enhances the need to optimize congestion management in order to reduce the resulting energy losses. Congestion forecasts can give the grid operator valuable time for countermeasures, and in this way supports optimization of congestion management. In addition, the recently introduced 'Redispatch 2.0' regulation in Germany makes forecasts of curtailments also necessary for distribution system operators. This paper addresses the need for congestion predictions by presenting three ANNs predicting vertical power flow on MV/HV transformers, wind generation connected to HV grid, and component loading for n-1 security cases. This enables potential congestion to be detected in two ways: by predicting the grid injections, and by forecasting component loads. Validation, which was carried out on the basis of actual data provided by the grid operator, showed that this could be promising approaches for grid operators. The comparison of the approaches reveals that in case for vertical power prediction, the presented ANN method is more promising than the parametric approach. Future work has to take possible uncertainties of the predictions into account. Further, the determination of which RE plant has to be curtailed should also be integrated.

ACKNOWLEDGMENT

The authors would like to thank Deutscher Wetterdienst for wind speed analysis data of the COSMO-DE Model and all project partner of the research project "enera." They also thank EWE NETZ GmbH for providing time series of aggregated load.

REFERENCES

[1] H. K. Jacobsen and S. T. Schröder, "Curtailment of renewable generation: Economic optimality and incentives," *Energy Policy*, vol. 49, pp. 663–675, Oct. 2012.
[2] X. Li and Q. Xia, "Stochastic optimal power flow with network reconfiguration: Congestion management and facilitating grid integration of renewables," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo. (T&D)*, Oct. 2020, pp. 1–5.
[3] "Quartalsbericht Netz-und Systemsicherheit-Gesamtes Jahr 2020," Bundesnetzagentur Für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen, Bundesnetzagentur, Bonn, Germany, Tech. Rep., 2021.
[4] A. Hoffrichter, K. Kollenda, M. Schneider, and R. Puffer, "Simulation of curative congestion management in large-scale transmission grids," in *Proc. 54th Int. Universities Power Eng. Conf. (UPEC)*, Sep. 2019, pp. 1–6.
[5] P. Staudt, B. Rausch, J. Garttner, and C. Weinhardt, "Predicting transmission line congestion in energy systems with a high share of renewables," in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–6.

[6] "Leitfaden zum EEG-Einspeisemanagement Abschaltrangfolge, Berechnung von Entschädigungszahlungen und Auswirkungen auf die Netzentgelte," Bundesnetzagentur Für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen, Bundesnetzagentur, Bonn, Germany, Tech. Rep., 2014.
[7] *Gesetz Zur Beschleunigung des Energieleitungsausbaus*, Bundesgesetzblatt Jahrgang, Germany, May 2019, no. 19.
[8] *Redispatch Als Teil Des Marktlichen Engpassmanagements*, B. B. der Energie-und Wasserwirtschaft, Berlin, Germany, Aug. 2015.
[9] K. Brauns, C. Scholz, A. Schultz, A. Baier, and D. Jost, "Vertical power flow forecast with LSTMs using regular training update strategies," *Energy AI*, vol. 8, May 2022, Art. no. 100143.
[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[11] S. Ilić, A. Selakov, S. Vukmirović, A. Erdeljan, and F. Kulić, "Short-term load forecasting in large scale electrical utility using artificial neural network," *J. Sci. Ind. Res.*, vol. 72, no. 12, pp. 739–745, 2013.
[12] P. Staudt, B. Rausch, Y. Träris, and C. Weinhardt, "Predicting redispatch in the German electricity market using information systems based on machine learning," in *Proc. Int. Conf. Inf. Syst.*, Dec. 2018, pp. 1–17.
[13] E. Memmel, D. Peters, R. Völker, F. Schuldt, K. Maydell, and C. Agert, "Simulation of vertical power flow at MV/HV transformers for quantification of curtailed renewable power," *IET Renew. Power Gener.*, vol. 13, no. 16, pp. 3071–3079, Dec. 2019.
[14] E. Memmel, S. Schluters, R. Volker, F. Schuldt, K. Von Maydell, and C. Agert, "Forecast of renewable curtailment in distribution grids considering uncertainties," *IEEE Access*, vol. 9, pp. 60828–60840, 2021.
[15] J. P. S. Catalao, H. M. I. Pousinho, and V. M. F. Mendes, "An artificial neural network approach for short-term wind power forecasting in Portugal," in *Proc. 15th Int. Conf. Intell. Syst. Appl. Power Syst.*, Nov. 2009, pp. 1–5.
[16] K. Bhaskar and S. N. Singh, "AWNN-assisted wind power forecasting using feed-forward neural network," *IEEE Trans. Sustain. Energy*, vol. 3, no. 2, pp. 306–315, Apr. 2012.
[17] Z. Lin and X. Liu, "Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network," *Energy*, vol. 201, Jun. 2020, Art. no. 117693.
[18] A. Srivastava, D. Steen, L. A. Tuan, O. Carlson, I. Bouloumpasis, Q. Tran, and L. Lemius, "Development of a DSO support tool for congestion forecast," *IET Gener., Transmiss. Distrib.*, vol. 15, no. 23, pp. 3345–3359, Dec. 2021.
[19] R. Fainti, M. Alamaniotis, and L. H. Tsoukalas, "Three-phase congestion prediction utilizing artificial neural networks," in *Proc. 7th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Jul. 2016, pp. 1–5.
[20] M. Alali, F. N. Shimim, Z. Shahooei, and M. Bahramipanah, "Intelligent line congestion prognosis in active distribution system using artificial neural network," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2021, pp. 1–5.
[21] D. Peters, W. Heitkoetter, R. Völker, A. Möller, T. Gross, B. Petters, F. Schuldt, and K. V. Maydell, "Validation of an open source high voltage grid model for AC load flow calculations in a delimited region," *IET Gener., Transmiss. Distrib.*, vol. 14, no. 24, pp. 5870–5876, Jun. 2020.
[22] H. Schermeyer, C. Vergara, and W. Fichtner, "Renewable energy curtailment: A case study on today's and tomorrow's congestion management," *Energy Policy*, vol. 112, pp. 427–436, Jan. 2018.
[23] P. Wiest, K. Frey, K. Rudion, and A. Probst, "Dynamic curtailment method for renewable energy sources in distribution grid planning," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Jul. 2016, pp. 1–5.
[24] L. Michi, M. Migliori, A. C. Bugliari, B. Aluisio, G. M. Giannuzzi, and E. M. Carlini, "Transmission network expansion planning: Towards enhanced renewable integration," in *Proc. AEIT Int. Annu. Conf.*, Oct. 2018, pp. 1–5.
[25] S. A. Kalogirou, "Artificial neural networks in renewable energy systems applications: A review," *Renew. Sustain. Energy Rev.*, vol. 5, no. 4, pp. 373–401, Dec. 2001.
[26] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, "Data mining: Practical machine learning tools and techniques," in *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. San Mateo, CA, USA: Morgan Kaufmann, 2016, pp. 1–621.
[27] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometrics Intell. Lab. Syst.*, vol. 39, no. 1, pp. 43–62, 1997.
[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[29] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)-Conf. Track*, 2015.
[30] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 2623–2631.
[31] Y. Ozaki, Y. Tanigaki, S. Watanabe, and M. Onishi, "Multiobjective tree-structured Parzen estimator for computationally expensive optimization problems," in *Proc. Genetic Evol. Comput. Conf.*, Jun. 2020, pp. 533–541.
[32] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, and V. A. Kamaev, "A survey of forecast error measures," *World Appl. Sci. J.*, vol. 24, no. 24, pp. 171–176, 2013.
[33] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and f-score, with implication for evaluation," in *Advances in Information Retrieval*, D. E. Losada and J. M. Fernández-Luna, Eds. Berlin, Germany: Springer, 2005, pp. 345–359.
[34] M. Abadi, "Tensorflow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Design Implement., (OSDI)*, 2016, pp. 265–283.
[35] W. McKinney, "Data structures for statistical computing in Python," in *Proc. 9th Python Sci. Conf.*, S. van der Walt and J. Millman, Eds., 2010, pp. 56–61.
[36] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, and R. Kern, "Array programming with numpy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.
[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 10, pp. 2825–2830, Jul. 2017.
[38] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with Python," in *Proc. Python Sci. Conf.*, 2010, Art. no. 25080.
[39] T. G. Smith. (2017). *Pmdarima: Arima Estimators for Python*. Accessed: Nov. 10, 2022. [Online]. Available: http://www.alkaline-ml.com/pmdarima
[40] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 competition: 100,000 time series and 61 forecasting methods," *Int. J. Forecasting*, vol. 36, no. 1, pp. 54–74, Jan. 2020.
[41] M. Baldauf, J. Förstner, S. Klink, T. Reinhardt, C. Schraff, A. Seifert, and K. Stephan, *Kurze Beschreibung des Lokal-Modells Kürzestfrist COSMO-DE (LMK) und Seiner Datenbanken Auf Dem Datenserver des DWD*. Offenbach, Germany: Deutscher Wetterdienst, 2016.
[42] A. Hammer, J. Kühnert, K. Weinreich, and E. Lorenz, "Short-term forecasting of surface solar irradiance based on Meteosat-SEVIRI data using a nighttime cloud index," *Remote Sens.*, vol. 7, no. 7, pp. 9070–9090, 2015.
[43] H. Xiong, "Unbalanced data set classification based on convolutional neural network," in *Proc. Int. Conf. Comput. Netw., Electron. Autom. (ICCNEA)*, Sep. 2021, pp. 186–190.

ELENA MEMMEL received the B.Sc. degree in geoecology from the Technical University of Freiberg, Germany, in 2014, and the M.Sc. degree in environmental modeling from the University of Oldenburg, Germany, in 2017. She is currently pursuing the Ph.D. degree with the Department of Energy Systems Technology, DLR Institute of Networked Energy Systems, Oldenburg, Germany. Her research interests include renewable energy integration, modeling and simulation of power system operation, and machine learning techniques.

THOMAS STEENS received the master's degree in environmental engineering from The Technical University of Braunschweig, Germany, in 2019. From 2019 to 2022, he was a Researcher with the Department of Energy Systems Technology, DLR Institute of Networked Energy Systems, with a focus on energy management. His research interest includes the usage of machine learning techniques to support energy management systems.

SUNKE SCHLÜTERS received the Ph.D. degree in pure mathematics from the Carl von Ossietzky University of Oldenburg, Germany, in 2015, and switched to the field of energy management research, in 2018. He is currently a Researcher and a Project Manager on energy management strategies and algorithms with the DLR Institute of Networked Energy Systems. His research interest includes machine learning-based self-optimizing energy management systems.

RASMUS VÖLKER received the degree in electrical engineering from the Kiel University of Applied Sciences, Germany, and the master's degree in the specialization of electrical energy technology, in 2015. From 2015 to 2022, he was a Project Manager with the Department of Energy Systems Technology, DLR Institute of Networked Energy Systems, Oldenburg, Germany. His main research interests include reliable power grids with a high share of flexible and decentralized generation.

FRANK SCHULDT received the Diploma degree in electrical engineering from the University of Applied Sciences Emden, Germany, in 1994. From 1994 to 2011, he held various industrial positions. He is currently the Deputy Head of the Department of Energy Systems Technology and the Head of the Flexibility Options and Grid Services Group, DLR Institute of Networked Energy Systems, Oldenburg, Germany. His research interests include integration of flexibility options in energy systems, reliability, availability, maintainability, and safety in power grids, and the design of system services for energy systems with a high share of power electronic-based generation.

KARSTEN VON MAYDELL received the B.A. and M.Sc. degrees in physics from the University of Oldenburg, Germany, in 1997 and 2000, respectively, and the Ph.D. degree in physics from the University of Marburg, Germany, in 2003. From 2000 to 2006, he worked as a Graduate Research Assistant, a Postdoctoral Researcher, and a Project Manager with Helmholtz Zentrum Berlin. From 2006 to 2007, he worked as a Project Manager of research and development with Q-Cells AG, Thalheim, Germany. From 2007 to 2008, he was a Group Leader with the Energy and Semiconductor Research Laboratory, University of Oldenburg. He was the Head of the Division Photovoltaics, from 2008 to 2014, and the Head of the Division of Energy Systems and Storage, NEXT ENERGY Research Institute, Oldenburg, Germany, from 2014 to 2017. Since 2017, he has been the Head of the Department of Energy Systems Technologies, DLR Institute of Networked Energy Systems. His research interests include the design of energy systems, smart energy management for grid-connected and off-grid energy systems, integration of flexibilities in energy systems, robust operation of power grids, and power electronics design.