

Received October 18, 2021, accepted December 1, 2021, date of publication December 10, 2021, date of current version December 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3134373

Static Grid Equivalent Models Based on Artificial Neural Networks

ZHENG LIU¹, JAN-HENDRIK MENKE^{1,2}, NILS BORNHORST¹,
AND MARTIN BRAUN^{1,3}, (Senior Member, IEEE)

¹Department of Energy Management and Power System Operation, University of Kassel, 34121 Kassel, Germany

²Amprion GmbH, 44263 Dortmund, Germany

³Fraunhofer Institute for Energy Economics and Energy System Technology, 34119 Kassel, Germany

Corresponding author: Zheng Liu (zheng.liu@uni-kassel.de)

This work was supported by the University of Kassel, Kassel, Germany, through the funds for open access publications.

ABSTRACT Power systems are rapidly and significantly changing due to the increasing penetration of distributed energy resources (DERs) and the rapid growth of widespread grid interconnections. An increasing number of grid operators is thus interested in the reduced equivalent representation of a large, interconnected power system to reduce the amount of required computational resources and data exchange, e.g., between grid operators. However, state-of-the-art grid equivalents become more and more inapplicable since they are analytically calculated for one specific grid state. They cannot properly be adapted to grid state changes and the behavior of the increasingly used controllers, such as reactive power controllers of DERs. Therefore, an innovative grid equivalent based on artificial neural networks (ANN) is proposed which overcomes the drawbacks of the state-of-the-art grid equivalents as follows: 1) Using supervised ANNs with feedforward and recurrent architectures, power systems can be equivalently represented adaptively and thus more accurately. 2) A feature selection method identifies the elements in the grid with high sensitivity on the boundary enabling a reduction of grid data required for the ANN-based equivalent. 3) To guarantee data confidentiality and cybersecurity, an additional unsupervised ANN, an Autoencoder, is used for obfuscation of the data which is required to be exchanged among grid operators, while the relevant information of the original data is preserved, maintaining the estimation accuracy. The ANN-based approach is analyzed and evaluated with two German benchmark grids and representative scenarios. The results demonstrate that the proposed approach outperforms the state-of-the-art radial equivalent independent method.

INDEX TERMS Grid equivalent, feedforward neural networks, recurrent neural networks, autoencoder.

I. INTRODUCTION

A. MOTIVATION

In 2014, the European Commission set a new target of 15% electricity interconnection by 2030, i.e., all Member States should achieve a level of electricity interconnection (electricity transport across borders to neighboring countries) to at least 15% of their installed generation capacity [1]. Power systems worldwide increase in size and complexity due to the increasing penetration of distributed energy resources and the rapid growth of widespread interconnections. Grid operators have to cooperate closely at their borders to neighbors on the same or to other voltage levels. Due to the

The associate editor coordinating the review of this manuscript and approving it for publication was Bin Zhou¹.

interconnection, there are interdependencies that need to be considered between neighbor grid operators. Special attention has to be given to the modeling and simulation of a large and interconnected power grid. A standard solution is to use the equivalent model in place of a complete detailed model of the neighbor grids. There can be various reasons for using an equivalent model. The main reasons are the following:

1. Practical limitations on the computational resources for power market behavioral analysis, grid monitoring, operational and planning studies of large power systems [2]–[5].
2. As the electrical distance from the point of interest increases, the required detail of modeling of the remote location is lower [3], [6].
3. An increasing number of grid operators are interested in cross-network cooperation, such as horizontal cooperation

(among transmission system operators (TSOs) and distribution system operators (DSOs)) and vertical cooperation (TSO-DSO) [7], [8]. The involved grid operators are usually unwilling to share their grid data because of confidentiality and security-relevant reasons [1], [4], [9].

Therefore, obtaining an equivalent part of the power system that is analyzed is of great importance. According to the representation of the model and its intended use, grid equivalent techniques can be broadly classified into static and dynamic [10]. In this paper, the term “grid equivalent” only refers to the static equivalent that is used for the analysis of quasi-steady states, such as grid operation, planning, and market-oriented studies. Fig. 1 shows the general case of an interconnected power system, which is divided into the internal subsystem (IS) and the external subsystem (ES) (IS and ES could belong to the same or different grid operators). The former, in which engineers are interested, remains unmodified, and a simple grid equivalent model represents the latter by using grid reduction methods. The reduced interconnected power system should represent the original system as accurately as possible.

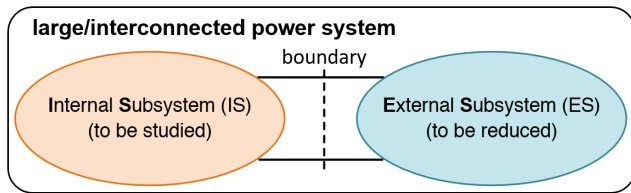


FIGURE 1. Illustration of subsystems including a grid equivalent.

B. LITERATURE REVIEW ON STATIC EQUIVALENTS

Many different grid equivalent methods have been developed over the last decades. The most classical grid equivalents are the Ward equivalent and the radial equivalent independent (REI) equivalent.

The Ward equivalent was initially proposed in [11] and then further discussed in [12]–[14]. It disregards all of the buses in the ES and models the ES as a set of equivalent impedance, shunts, and power injections attached to the boundary buses via the Gaussian elimination. To approximately consider the reactive power response of the ES to the IS, the extended Ward (xWard) equivalent was proposed where a fictitious PV bus is added without power injection at each boundary bus (see Fig. 2 a) and c)). The significant limitation of the Ward equivalent is that the operating points of the assets in the ES cannot be simply adapted when they change [11], [15]. Therefore, attempts have been made to minimize load flow errors caused by operating point change; one of them is called radial equivalent independent method (REI). The idea behind the REI equivalent is to shift the power injection in the ES to one or more fictitious REI buses [16]. Afterward, the passive external grid area resulting from the power shift can be reduced by the Gaussian elimination, and the original power injection is substituted by equivalent power injections at the REI buses attached to the boundary

buses (see Fig. 2 a) and b)). An adaptation to other operating points is possible without repeating the equivalent calculation process, with the help of a simple scaling of the operating points of the equivalent devices [10].

In addition, several researchers have used bus aggregation techniques to reduce large power networks. E.g., approximate equivalent networks have been generated based on power transfer distribution factor (PTDF) matrices [17], [18], where buses with a similar contribution to a designated internal subsystem are grouped in zones. Each of these zones is aggregated to a single bus. Another bus aggregation approach is based on the local marginal prices (LMPs) [19], [20]. Clusters of buses with almost identical LMPs are reduced using the REI equivalent. The application of this kind of equivalents is intended mainly for transmission cost allocations and electricity markets.

An obtained equivalent grid based on the methods mentioned above is generally only accurate for a specific grid state. Larger deviations from this grid state lead to less precision [15], [21]. In the intelligent grid paradigm, the switching statuses can be changed for congestion management [22], and different controllers, e.g., Q(V) control for distributed energy resources (DERs), are widely used and play a key role in improving the grid stability by adjusting operating points. These complicated interactions are difficult to be accurately represented and render the use of grid equivalents with acceptable accuracy more challenging. Also, the strong fluctuations of DERs lead to frequent grid state changes. The fundamental solution to this problem is to repeat calculating the grid equivalent after each grid state change. However, frequently recalculating the grid equivalent is usually not applicable in practice due to the high requirement of grid data, the high computational complexity, and the fact that this would require an increased data exchange, e.g., between adjacent grid operators [15].

C. CONTRIBUTION AND ORGANIZATION

Inspired by the application of an artificial neural network (ANN) for state estimation in [23], an innovative adaptive grid equivalent method based on ANN with feedforward

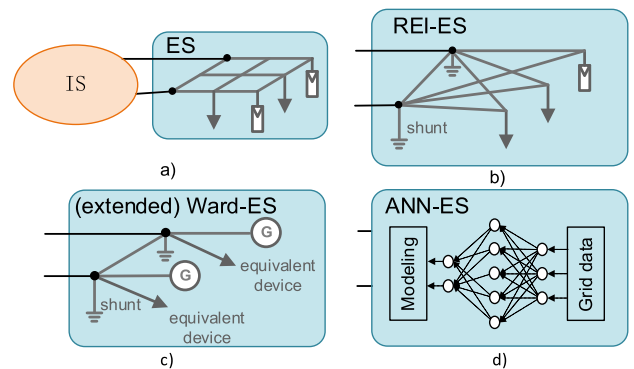


FIGURE 2. Scheme of grid equivalents. a) original power grid; b) REI equivalent grid; c) (extended) ward equivalent grid; d) ANN-based equivalent.

architecture has been proposed in our previous works [15], [24]. The method utilizes ANN to approximate a power grid through extensive training without using the original grid model. After the ANN model has been trained to sufficient accuracy, it yields as outputs the power flows between the IS and the ES at the boundary buses at a very low computational complexity and a high accuracy (see Fig. 2 a) and d)). Input to the ANN is the current grid parameter vector containing, e.g., the operating points of the loads and the DERs and the switching states. However, the previous approach has some limitations: 1) The high amount of grid data that was required to train the ANN, and some of the grid data had only a small impact on the accuracy; 2) Considering grid data confidentiality and cybersecurity, exchanging grid data may not be acceptable when different grid operators are involved; 3) Only feedforward architecture was used. Other kinds of ANN architectures have the potential to increase the estimation accuracy; 4) Lack of automatization and generalization.

In this paper, this approach is further developed by implementing the following improvements and extensions.

1. A component *feature selection* is implemented such that grid elements (and the corresponding grid data) with a high impact on the grid equivalent can be identified. By considering only the identified (featured) grid elements, the amount of grid data required to train the ANN can be reduced.
2. An ANN-based *data obfuscation* (Autoencoder) is proposed to obfuscate the grid data for confidentiality reasons. The obfuscated data rather than the original data is then used as input for the subsequent grid equivalent ANN. In the data obfuscation process, the data properties required for the grid equivalent ANN to estimate the interactions at the boundary buses are maintained, while the original data cannot be reconstructed from the obfuscated one without knowledge of the data obfuscation ANN.
3. An ANN with recurrent architecture is considered for the grid equivalent. Its feedback connections are more appropriate in power grids where feedback is involved, e.g., in DERs with local voltage control.

These developments are implemented as a module for grid equivalents with automated parametrization, training, simulation, and evaluation. Comprehensive simulations are carried out, considering horizontal and vertical equivalent scenarios, operating points of grid assets, voltage fluctuation, switching statuses of lines, tap changer statuses, reactive power controllers, and measurement errors. To validate the advantages of the ANN approach, its performance is compared to that of the previous ANN-based approach and the state-of-the-art REI equivalent. Due to the above-mentioned limitations, the REI equivalent could be shown to outperform the Ward and xWard equivalent in our preliminary work [15]. Therefore, in this paper, only the performance of the proposed equivalent to the REI equivalent (not to the Ward and xWard equivalent) is compared.

The structure of this paper is arranged as follows: In Section II, together with introductions of theoretical

fundamentals, the proposed ANN-based scheme is described. In Section III, the test grids and scenarios are presented. The evaluation with different grids and scenarios is performed in Section IV. Finally, in Section V, conclusions are drawn.

II. GRID EQUIVALENT USING ARTIFICIAL NEURAL NETWORKS

The most desirable property of a grid equivalent method is that it should represent the effect of the ES on the interconnection between the IS and the ES (i.e., on the boundary buses) as accurately as possible. This effect depends on the current grid state of the overall grid (IS + ES). Therefore, the goal of the grid equivalent is, in effect, to adaptively estimate the relationship between the grid state (inputs) and the interactions at the interconnection (outputs). This goal is represented by the core component *interaction estimation* of the proposed approach, see Fig. 3 a) middle. Starting with this component, the proposed scheme for the ANN-based grid equivalent is described successively.

A. ANN-CONFIGURATION FOR INTERACTION ESTIMATION

An artificial neural network imitates a biological neural network and can discover patterns adaptively in data sets because of its flexibility and nonlinearity. It has been used to learn from experiences complex input-output relationships with high accuracy [33]. The training of an ANN-based grid equivalent is realized within the component *interaction estimation*. As Fig. 3 illustrates, it receives the training input data (grid data set of historical grid states) and training output data (interactions, e.g., power exchange at the interconnection) and trains an ANN-estimator such that the outputs of the ANN approximate the given output training data as close as possible. In this component, two supervised ANNs with different architecture are considered.

1) FEEDFORWARD ARCHITECTURE

The first supervised ANN is a feedforward neural network (FNN). As its name implies, FNN is a framework having only direct connections between neurons of one layer and those of the next layer as the ANN depicted in Fig. 2 d). The input vector \mathbf{x} of each neuron is multiplied with the vector of weights \mathbf{W} and the resulting scalar is passed through an activation function f to produce the output $\mathbf{y} = f(\mathbf{W}^T \mathbf{x})$ [32]. In many cases, the output also includes a bias. For the regression task in this paper, a multi-layer FNN is used.

2) RECURRENT ARCHITECTURE

Unlike in FNNs, in recurrent neural networks (RNNs), feedback connections exist from the output of neurons to the inputs of neurons (either their own inputs or the inputs of other neurons). RNN is a dynamic system, and it can use its internal control state \mathbf{h}_t to process sequences of inputs [35], [36]. The state \mathbf{h}_t not only depends on the current inputs, but also on the previous state \mathbf{h}_{t-1} , see Fig. 4 a). Accordingly, the general mathematical representation of a

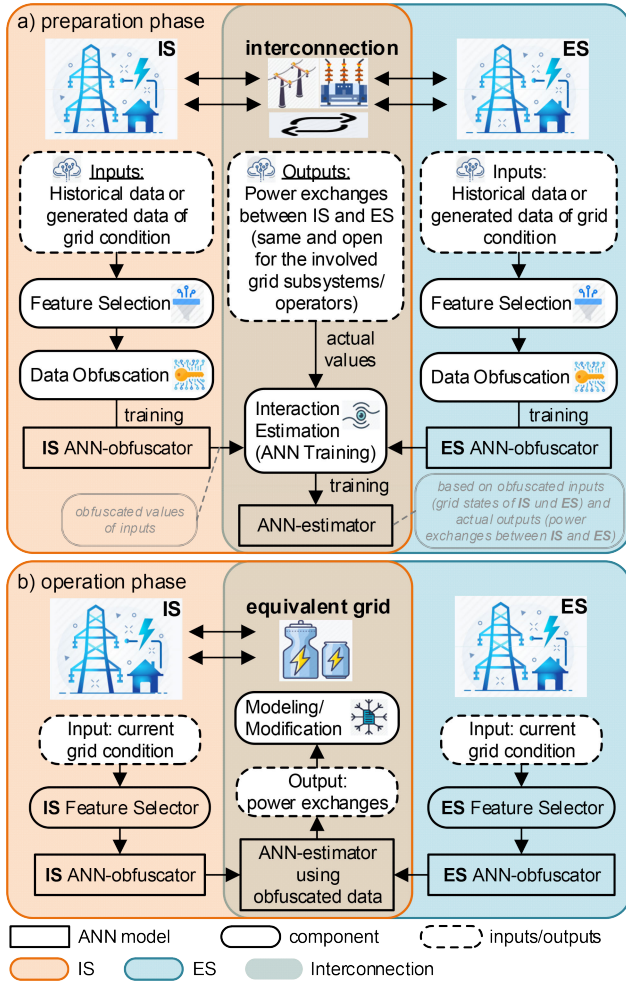


FIGURE 3. Scheme for the ANN-based grid equivalent for two grid operators. a) Preparation phase; b) Operation phase.

RNN for one pattern is presented by (1) and (2):

$$h_t = f_h \left(W_{xh}^T \cdot x_t + W_{hh}^T \cdot h_{t-1} \right) \quad (1)$$

$$y = f_o \left(W_{hy}^T \cdot h_t \right) \quad (2)$$

W_{xh}^T : transposed matrix of weights connecting the input x_t to the state h_t

W_{hh}^T : transposed matrix of weights connecting the previous state h_{t-1} to the current state h_t

W_{hy}^T : transposed matrix of weights connecting the input h_t to the output y

The output of the current input depends on the past computation such that the ANN exhibits a memory [36].

In power systems, many recursive processes exist, e.g., controllers such as local voltage controllers. E.g., in a Q(V)-control, the reactive power Q fed-in by a DER depends on the local voltage V which depends on Q. When the local voltage V changes, Q changes as well, which changes Q again, forming a feedback loop. To represent the recursive feedback processes, the recurrent architecture is modified, i.e., instead of time series $[x_t, x_{t+1}, x_{t+2}]$, static and duplicated grid state of a single time-step $[x_t, x_t, x_t]$ is

considered as input data, see Fig. 4 b). The output y_t of each RNN-layer is obtained after two or more recursive iterations.

From many varieties of RNNs, the long short-term memory (LSTM) architecture has been used in a wide range of applications because of its ability to overcome the gradient vanishing problem [35], [37]. Thus, the LSTM-RNN for the ANN-based grid equivalent has been implemented.

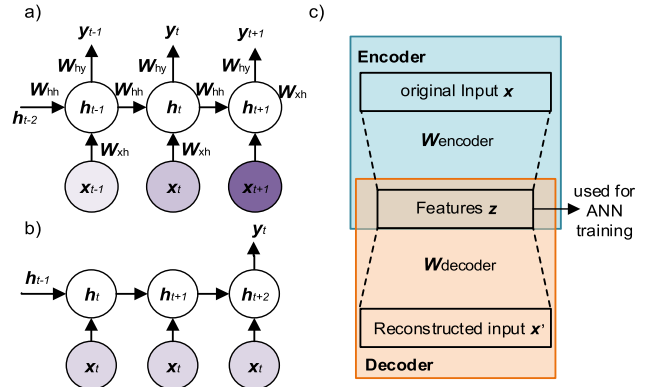


FIGURE 4. a) RNN architecture in a time-sequential manner; b) modification of RNN architecture for grid equivalent; c) scheme for Autoencoder.

B. TRAINING DATA PRE-PROCESSING

1) FEATURE SELECTION

The observed grid elements (see Table 1) for the ANN-based equivalent have different impacts on the interconnections due to various capacities and distances to the boundary. The elements with a small impact on the interaction require the same amount of grid data for the ANN training as those with a high impact. However, they improve the estimation accuracy only marginally. Thus, the component *feature selection* is proposed, where those elements with a high impact on the interconnections, which are the critical (featured) elements, are identified based on a set of sensitivity analyses. In general, the sensitivity analysis is based on a single static grid state. To find the critical elements associated with power injection, a sensitivity analysis for each time step (grid state) of the used grid data set is carried out. The analysis yields the sensitivity of a power injection to the power exchange at the interconnection $\left(\frac{\Delta P_{interaction}}{\Delta P_{element}} \right)$ and $\left(\frac{\Delta Q_{interaction}}{\Delta Q_{element}} \right)$ and to the voltage magnitude at the boundary buses $\left(\frac{\Delta V_{boundary}}{\Delta P_{element}} \right)$ and $\left(\frac{\Delta V_{boundary}}{\Delta Q_{element}} \right)$. This set of calculations is treated as the base case sensitivities. A DER or load is critical for the grid equivalent if any corresponding result for any time step meets the predefined thresholds, cf. Table 2. To determine the sensitivities of the switching status of a single branch element (e.g., a line), a set of sensitivities is calculated for each switching status change (e.g., outages of every single line). The effect of the switching status change on the interconnection depends on its deviation to the corresponding base case sensitivities. The switching status of the branch element is considered for the ANN-based equivalent if its effect on the interaction meets the predefined thresholds

or any other branch elements become overloaded. The sensitivity analyses mentioned above are very computationally intensive. To deal with this problem, the fast parallel power flow calculation algorithms in [38] has been used, with which the computing time is substantially reduced.

TABLE 1. Observed elements and grid data in this paper.

Grid element	Related grid data
connection point	voltage magnitude
load	active & reactive power
DER	active & reactive power
line & switch	switching status
transformer	tap changer position
controller	control parameter

2) DATA OBFUSCATION

In practice, the IS and the ES could belong to the same or different grid operators. In the latter case, different grid operators are not willing to share their grid data with others. In view of this situation, it is essential and meaningful to obfuscate the grid data set before sharing. In the obfuscation process, the relevant information for the training of the ANN-estimator has to be maintained. Another ANN with unsupervised learning, Autoencoder (AE), is used. The unsupervised learning is supplied with unlabelled data sets (containing only the input data) and left to find properties in the data set and build a new model from it [34]. While supervised learning leads to regression and classification, an unsupervised-learning-based AE performs feature learning, clustering, obfuscation, and data dimensionality reduction [39], which is what the proposed method needs to obfuscate the grid data. As Fig. 4 c) illustrates, an AE consists of an encoder and a decoder. The encoder transforms the input \mathbf{x} and produces the obfuscated code \mathbf{z} :

$$\mathbf{z} = \mathbf{f}_{\text{encoder}}(\mathbf{W}_{\text{encoder}}^T \cdot \mathbf{x}) \quad (3)$$

The decoder then reconstructs the input data by obtaining outputs \mathbf{x}' that are as close as possible to the original input data \mathbf{x} :

$$\mathbf{x}' = \mathbf{f}_{\text{decoder}}(\mathbf{W}_{\text{decoder}}^T \cdot \mathbf{z}) \quad (4)$$

Different kinds of AEs aim to achieve different kinds of properties [40]. In this paper, an AE with stacked architecture is used, which is most appropriate for cybersecurity applications [39], is chosen. The stacked AE trained is symmetrical, and both the encoder and the decoder are fully connected FNNs. The middle-most layer, the output of the encoder \mathbf{z} (cf. Fig. 4 c)), is used as the obfuscated data for data exchange and the subsequent training of the ANN-estimator. The original input data \mathbf{x} can only be reconstructed by the decoder that is simultaneously trained in the same training process, i.e., by an FNN with $\mathbf{W}_{\text{decoder}}$. Since the latter is not shared, it is unknown to the other grid operator, and a reconstruction of the original input data would only be possible for the grid operator delivering the obfuscated data \mathbf{z} .

The data obfuscator is only used for the input data. The output data set, which is the data at the interconnection, is generally the same and open for both involved grid operators (cf. Fig. 3 a) and can be directly used without obfuscation.

It is assumed that an RNN-estimator based on obfuscated data is used for a grid equivalent task. The original grid data for IS and ES (\mathbf{x}_{IS} and \mathbf{x}_{ES}) are separately obfuscated by (5) and (6). The vector concatenation \mathbf{x}_t of their results \mathbf{z}_{IS} and \mathbf{z}_{ES} (equation (7)) is the input for the RNN that estimates the interaction \mathbf{y} at the interconnection, see equations (1), (2), and (8).

$$\mathbf{z}_{\text{IS}} = \mathbf{f}_{\text{encoder}}(\mathbf{W}_{\text{IS-encoder}}^T \cdot \mathbf{x}_{\text{IS}}) \quad (5)$$

$$\mathbf{z}_{\text{ES}} = \mathbf{f}_{\text{encoder}}(\mathbf{W}_{\text{ES-encoder}}^T \cdot \mathbf{x}_{\text{ES}}) \quad (6)$$

$$\mathbf{x}_t = [\mathbf{z}_{\text{IS}} \ \mathbf{z}_{\text{ES}}] \quad (7)$$

$$\mathbf{y} = \mathbf{f}_o(\mathbf{W}_{\text{hy}}^T \cdot \mathbf{f}_h(\mathbf{W}_{\text{xh}}^T \cdot \mathbf{x}_t + \mathbf{W}_{\text{hh}}^T \cdot \mathbf{h}_{t-1})) \quad (8)$$

C. MODELING AND MODIFICATION

Based on the selected and obfuscated grid data, through proper training, an ANN can be used to find a mapping from different grid states to the interactions at the interconnection. In our proposed approach, the estimated interactions are modeled as equivalent devices attached to the boundary in the IS. Which kind of equivalent devices to be implemented depends on the data type the estimated interaction contains, i.e., the estimated power values are represented by equivalent adaptive loads, and the estimated voltage values are modeled as equivalent adaptive generators or external grids, such that the effect of the ES is approximated. As shown in Fig. 3 b), grid operators exchange their selected and obfuscated grid data in the operation phase. The corresponding interactions estimated by the ANN-estimator are implemented by modifying the equivalent devices at the boundary, e.g., by modifying the values of the equivalent adaptive loads.

III. TEST POWER GRIDS AND SCENARIOS

To better understand the functionalities of the proposed approach, in this section, the ANN-based grid equivalent is applied and investigated based on two German benchmark grids and two scenarios, i.e., as a horizontal equivalent and as a vertical equivalent, see Table 2.

The motivation for the horizontal equivalent is that within one grid (and one voltage level), the requirements for detailed modeling of distant grid areas decreases, and for the calculation with large grids increases (cf. the first two reasons for using a grid equivalent in Section I-A). The related grid areas are represented by using the proposed approach. The grid is a slightly modified version of the medium voltage grid Oberrhein in [30]. The configurations for the ANN-based approach are listed in Table 2. The training inputs consist of voltages at the boundary buses, the operating points of loads and DERs, the switching statuses of lines, and the individual controller parameters. They are all varying over time within the time series. The outputs are the corresponding power exchanges at the boundary buses. The component

feature selection is activated and deactivated, respectively, investigating its effect on the accuracy of estimation. As no grid data exchange between different operators is involved, the component *data obfuscation* is deactivated. Both FNNs and RNNs are considered. Within this scenario, the following analyses are performed: 1. effects of the feature selection on the accuracy; 2. comparison of different ANN-architectures; 3. comparison of the proposed method with conventional methods.

Grid equivalents are also required in cross-grid studies where different grid operators are involved. Thus, a modified SimBench grid [41] with multi-voltage levels and operators is used for the vertical scenario. The extra-high-voltage (EHV) level containing eight connection points operated by the TSO and the high-voltage (HV) level operated by the DSO are interconnected via three transformers. Except for the voltages at the connection points, the operating points, the switching statuses of lines, and time-varying tap-changer positions are considered for this scenario (see Table 2). To train an ANN-estimator to adaptively represent the interactions that are influenced by both grid operators, TSO and DSO have to exchange their grid data sets. Hence, the component *data obfuscation* is activated. The following analyses are performed: 1. effects of different configurations on data obfuscation; 2. effects of obfuscated data on accuracy; 3. comparison of the proposed approach with conventional methods.

Also, horizontally interconnected grids can belong to different grid operators (e.g., DSOs’ or TSOs’ cooperation). The application of the proposed approach in the latter case is comparable to the scenario described above.


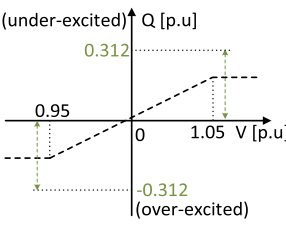
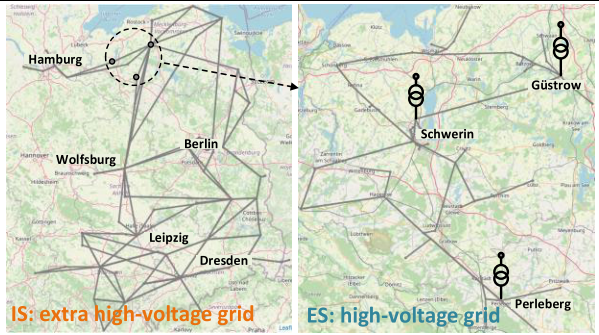
IV. SIMULATION

A. IMPLEMENTATION AND SIMULATION ASSUMPTION

The implementation of the proposed approach and the classical equivalent methods is done in the programming language Python, based on the grid analysis tool *pandapower* [30], the PyTorch package [28], [29], and the estimation tool in [23]. The implemented REI, Ward, and xWard methods are validated by comparing the results of our implementation with those provided by DIGSILENT PowerFactory [31]. Before and after the calculation of an equivalent, the static bus voltages deviate up to 10^{-6} p.u. for standard IEEE benchmark grids such as case9, case39, case118. *pandapower* is used as the tool for grid simulation.

For an ANN to accurately approximate the interaction at the interconnection for a given grid state, it is important to obtain a meaningful training data set, which covers as many grid situations as possible. An improved version of the scenario generator of our preliminary work [23] is used such that grid parameters for scenario generation are configurable through a simple string, i.e., in the framework of *pandapower*, the observed grid parameters follow

TABLE 2. Configuration and test power grids for different scenarios.

	Horizontal equivalent		Vertical equivalent
	Parameter	Value	Value
Configuration and information	Grid information	The medium voltage grid Oberrhein in [30] is located in the section of Rhine in the Upper Rhine Plain between Basel in Switzerland and Bingen in Germany. It has 179 buses, 51 DERs, 147 loads, and 1 connection point.	In the project SimBench [43], a benchmark dataset to support research in grid planning and operation has been developed. The SimBench grid used for the vertical equivalent consists of a EHV grid in northeast Germany and a HV grid for the region Schwerin, which belong to the control area of the German TSO 50Hertz. It contains 147 buses, 54 DERs, 52 loads, and 8 connection points.
	Inputs:	<ul style="list-style-type: none"> - voltage at the connection point - operating points of loads - operating points of DERs - switching statuses of lines - Q(V) controller parameters 	<ul style="list-style-type: none"> - voltage at the connection points - operating points of loads - operating points of DERs - switching statuses of lines - tap-changer positions
	Outputs	power interaction at the boundary	power interaction at the boundary (EHV/HV-interface)
	Feature selection	activated and deactivated	activated
	Sensitivity thresholds	$\frac{\Delta P_{interaction}}{\Delta P_{load/DER}} > 5 * 10^{-3}$; $\frac{\Delta Q_{interaction}}{\Delta Q_{load/DER}} > 5 * 10^{-3}$; $\frac{\Delta V_{boundary}}{\Delta P_{load/DER}} > 5 * 10^{-3}$; $\frac{\Delta V_{boundary}}{\Delta Q_{load/DER}} > 5 * 10^{-3}$	
	Data obfuscation	deactivated	activated and deactivated
	ANN architecture	FNN and RNN	FNN and RNN
Test power grid/system	 <p>50% DERs are controlled by the following droop curve individually. The max./min. Q are varied between 0 and 0.312/-0.312</p> 		 <p>IS: extra high-voltage grid ES: high-voltage grid</p>

their elements by dashes (–) and connected with each other by slashes (/), e.g., *load-p_mw/line-in_service/controller-cos_phi* means that the active power of load, the switching status of lines and the controller parameter $\cos \varphi$ are considered during the scenario generation.

For the horizontal equivalent scenario, an annual time-series (inputs) is generated with 15-min resolution for the input elements mentioned in Table 2. The corresponding annual interactions at the boundary (outputs) are obtained by power flow calculations. To simulate grid topology changes, a single random line is switched off at each time step. This generated data set is divided into two parts for training (first half) and grid simulation (second half). For the vertical equivalent scenario, the SimBench grid provides data sets of realistic time series for loads and DERs in 15-min over one year (35040 time steps). In addition, random tap changer positions and voltages of the eight external grids have been added for each time step (cf. Table 2). After an annual simulation, the corresponding power exchanges between the TSO and the DSO (outputs) are calculated. To simulate in a more realistic way, measurement errors are considered (maximum error of 1.5% for power measurements and maximum error of 0.5% for voltage measurements according to IEC 61869) and added to the input data set for the ANN. Furthermore, to let the ANN learn the periodicity of the realistic data set, a number (from 0 to 95) is added to the input data for each time step to represent the time of a day (24 hours per day, 4 time steps per hour \rightarrow 96 time steps per day). The data set from January to August is used for training and that from September to December for simulation. The complete TSO-DSO grid is only used for the grid data set generation using power flow calculations such that the outputs are correctly calculated according to the given inputs. In practice, it is not required to perform power flow calculations on the complete TSO-DSO grid as measurements of the boundary buses should be available for both grid operators.

To find a proper ANN, the optimization method in [25] is used to obtain the number of layers and other hyperparameters. *Adam* [26] is chosen to optimize the ANN's weights, and loss functions *L1Loss* [27] for ANN-estimator and *MSELoss* [28] for ANN-obfuscator are used. All results are produced with an Intel i7-4702MQ CPU (2.2 GHz), 16 GB of RAM (800 MHz), SSD storage, Python 3.7 on Ubuntu Linux with GPU (2*GeForce GTX 1080) acceleration.

In this paper, the accuracy of the ANN-estimator is assessed using the *Weighted Average Percentage Error* (WAPE):

$$\text{WAPE} = \frac{\sum_{i=1}^n |E_i - A_i|}{\sum_{i=1}^n |A_i|} * 100 \quad (9)$$

where: E_i – the estimated vector for time-step i ; A_i – the actual vector for time-step i ; n – time-step size of the observed data set. Compared to the most commonly used key performance indicator to measure estimation accuracy, the mean absolute percentage error, the WAPE overcomes the high sensitivity to outliers and the issues when the actual value is zero [42].

B. EVALUATION FOR HORIZONTAL EQUIVALENT

1) FEATURE SELECTION RESULTS

To find the critical elements for the interconnection, the component *features selection* is activated. With the predefined conditions of Table 2, the *feature selection* has found 273 critical elements. Compared to the original elements, the amount of required data set could be reduced by more than one-third (157 of 430), see Table 3. The original data set and the critical data set are used for the subsequent ANN training.

TABLE 3. Feature selection results for horizontal equivalent.

Grid parameter	Number before FS	Number after FS (reduced number)
V_{PCC}	1	1
$P_{load} \& Q_{load}$	147&147	84&84 (-126)
P_{DER}	51	46 (-5)
controller parameter	25	22 (-3)
S_{line}	59	36 (-23)
total	430	273 (-157)

2) ANN MODELS AND TRAINING RESULTS

According to our previous works [15], [24], an ANN-estimator with feedforward architecture, with the ID *FO* (see Table 4 for the definition of the IDs), is trained as a benchmark model based on the original grid data set (labeled with an *O* in its ID) and optimized hyper-parameters. To make a fair comparison, three other models with the IDs *ROI*, *FCI*, and *RCI* have been created, which inherit the hyper-parameters of *FO*, see Table 4. The differences between the FNN- and RNN-models are 1) Instead of the activation function ReLU used in the FNN models, the LSTM architectures represent the layers in the RNN models; 2) Each LSTM layer has a memory size of two, cf. Fig. 4. Models trained with critical data are labeled with a *C* in their IDs. To verify the effect of the critical data set and the component *feature selection*, two additional models, *FCO* and *RCI2*, are trained with re-optimized hyper-parameters, i.e., for the latter two models, the hyper-parameters are re-optimized with the reduced data set (the input size is 273) of only the critical data.

Fig. 5 illustrates the training results for different ANN models. The differences between the results of FNN (dark color) and that of RNN (light color) are obvious. In the diagrams for train loss, the RNN-models can almost always reach a lower loss with fewer epochs during the training process. The related training time (b) and estimation time for all the samples (d) for the RNN-models are higher than that for the FNN-models due to the recursive architecture of the RNN-models. With RNN, the outputs are more accurately estimated (c), and the accuracy is improved by up to 40%.

Compared to the *FO* and *ROI* (blue), with the same hyper-parameters, the ANN-models *FCI* and *RCI* based on critical data (orange) are successfully trained in less time due to the smaller input size. However, the estimation accuracy is slightly decreased. With the re-optimized hyper-parameters, the ANN-models *FCO* and *RCI2* (green) trained with the

TABLE 4. ANN-models for horizontal equivalent.

ID	Input size	Description
FO	430	FNN based on Original data and optimized hyper-parameter
ROI		RNN based on Original data and Inherited hyper-parameter of the model FO
FCI	273	FNN based on Critical data and Inherited hyper-parameter of the model FO
RCI		RNN based on Critical data and Inherited hyper-parameter of the model FO
FCO	273	FNN based on Critical data and Optimized hyper-parameter
RCI2		RNN based on Critical data and Inherited by hyper-parameter of the model FCO

TABLE 5. Hyperparameters for the ANN-models in Table 4.

Models	Hyperparameters
FO, (ROI), FCI, and (RCI)	Layer No./ Layer size/ Activation function
	1 / 530 / ReLU (LSTM)
	2 / 30 / ReLU (LSTM)
	3 / 370 / ReLU (LSTM)
FCO and (RCI2)	Layer No./ Layer size/ Activation function
	1 / 570 / ReLU (LSTM)
	2 / 530 / ReLU (LSTM)
	3 / 4 / linear

reduced data set even outperform those trained with the original data set, the FO and ROI (blue), in terms of estimation accuracy. The reasons for the longest training time for FCO and RCI2 are 1) During the re-optimization and the retraining, ANNs need more time to find an optimal relationship between the limited inputs (critical data set) and outputs; 2) The RNN-architecture is more complicate than the FNN-architecture so that its training time is longer.

To summarize, the results shown in Fig. 5 match the expectations: 1) The reduced and critical data set has nearly no effect on the accuracy, and, therefore, the requirement of data volume for data exchange in practice can be reduced substantially, cf. Table 3; 2) The implemented RNN-models improved the accuracy significantly.

3) INTERACTION EVALUATION

The best performing ANN-models FO, ROI, FCO, and RCI2 are applied in the operation phase in a time-series simulation. The estimated values are implemented as equivalent loads attached to the boundary buses and modified for each time step. Fig. 6 shows the deviation of the estimated active power and reactive with different equivalent models for the ANN-equivalents and for the repeated and non-repeated REI equivalent. REIrep means that the REI equivalent is recalculated in every time step to be able to adapt it to the changing grid conditions. Analogous to the results in Fig. 5 c), the interaction estimation is improved with the recurrent architecture (P-estimation improvement is more significant than the Q-estimation improvement). With the reduced grid data and re-optimized hyper-parameters, the accuracy for RCI2 (light green) is closest to that based on the original data (ROI

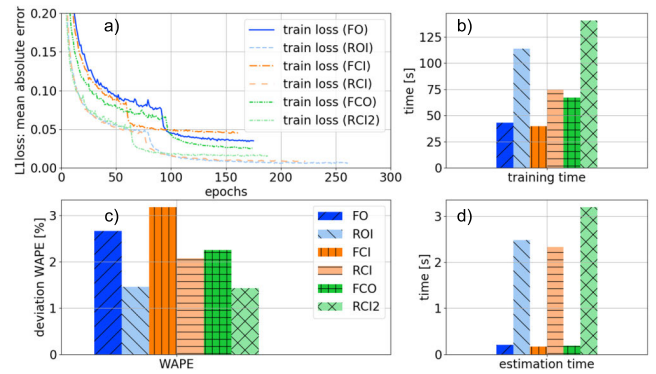


FIGURE 5. Results comparison for different ANN-models: a) train loss; b) training time; c) estimation accuracy measured in WAPE; d) estimation time for all samples.

(light blue)). The deviation for the conventional REI method (purple) without recalculation is about eight times larger than that of the ANN estimators. With the REIrep model, the interaction can be precisely calculated within 1.2s, see Table 6. However, the computation time for REI-calculation depends on the size and the complexity of the grid model and increases as the grid grows larger and more complex. Thus, frequent recalculation can require a long computation time which is usually not feasible. In addition, according to our experience, users often have to deal with a large sparse matrix occurring during the REI-calculation process. The sparsity can cause convergence problems in the calculation process. The most obvious drawback of the REIrep model is that grid operators have to provide comprehensive grid information (including a grid model) for each time step. Grid operators are reluctant to do this, and grid models for some grid fields are not even available in detail or lack updates, which can still cause deviations.

Considering these drawbacks for the REI-models, the proposed approach clearly outperforms the state-of-the-art REI approach since it yields almost instantaneous re-estimation and high estimation accuracy with only limited grid data and excluding grid models. Among all the ANN-models, the RNN-models ROI and RCI2 yield the best results. The RCI2 needs only reduced information of the current grid condition. The very short computation time enables an application in real-time.

C. EVALUATION FOR VERTICAL EQUIVALENT

1) FEATURE SELECTION RESULTS

The original data set for this scenario has a total of 359 parameters for 35040 (96*365) time steps, from which the feature selection has identified 166 critical parameters for the ANN-equivalent, see Table 7. Note that none of the single line outages has a significant impact on the interconnection between the TSO and the DSO due to the N-1 security criterion and the meshed grid topology. The observed loads and DERs are mostly residual devices to represent under-loaded medium voltage grids or high-power devices. Their large power fluctuation cannot be disregarded. Therefore, only a

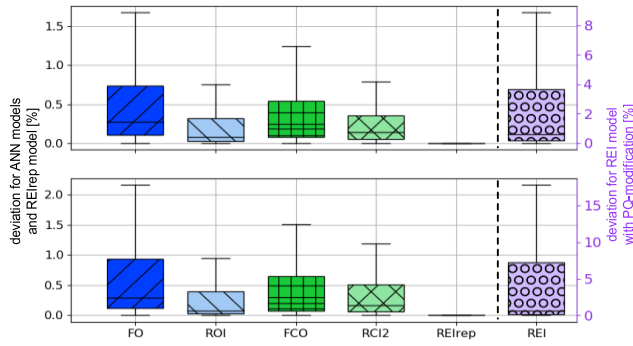


FIGURE 6. Power interaction deviation at the interconnection for different equivalent models.

TABLE 6. Mean computing time, accuracy (ACC), and required grid data for re-estimation or recalculation for different models.

ID	computing time [s]	1 st quartile to 3 rd quartile		required data
		P ACC [%]	Q ACC [%]	
FO	0.0015	0.05-0.75	0.70-0.92	data of current grid condition
ROI	0.0016	0.02-0.32	0.02-0.40	limited data of current grid condition
FCO	0.0015	0.06-0.55	0.05-0.60	limited data of current grid condition
RCI2	0.0017	0.02-0.35	0.02-0.48	limited data of current grid condition
REI	0	0.12-3.80	0.10-7.20	operating points
REIrep	1.14	0.00-0.00	0.00-0.00	all grid data and model

few loads and DERs are ignored. The deviations between using the original grid data and considering only the critical grid data are comparable to that for the horizontal scenario, cf. Fig. 5. Thus, for the remainder of this section - for the vertical scenario - only the results for the critical (reduced) data set are shown.

2) DATA OBFUSCATION EVALUATION

As explained in Section II-B-2), the outputs of the encoder are used as the obfuscated data. Also, the size of the encoder’s output layer is configurable, enabling obfuscated data with different dimensionality. A reduced dimensionality allows a data exchange with reduced data volume. The original data set is obfuscated with different reduction degrees (RD). Fig. 7 shows the results for an array consisting of 10 randomly selected active power operating points from the original data set. The obfuscated data with different RD and its related reconstructed data for one time step are shown in Fig. 7 a) and b). The obfuscated values deviate from the original data. With increasing RD, the dimensionality of the obfuscated data is reduced, and some information of the original data is lost due to the data compression (e.g., the deviation between blue points and black points in Fig. 7 b) are the largest). As a consequence, the reconstruction loss increases (see Fig. 7 b) and c) [40].

The general criteria for grid equivalents are “accurate”, “anonymous”, and “less data exchange”. In practice, under acceptable accuracy, grid operators can further reduce the data volume for data exchange with a nonzero RD. For the remainder of this section, the obfuscated data with RD0%

TABLE 7. Feature selection results for vertical equivalent.

Grid parameter	Number before FS	Number after FS (reduced number)
P_{load} & Q_{load}	52 & 52	48 & 48 (-8)
P_{DER}	54	49 (-5)
S_{line}	190	0 (-190)
POS_{ehv/hv_trafo}	3	3
$V_{external_grid}$	8	8
total	359	166 (-199)

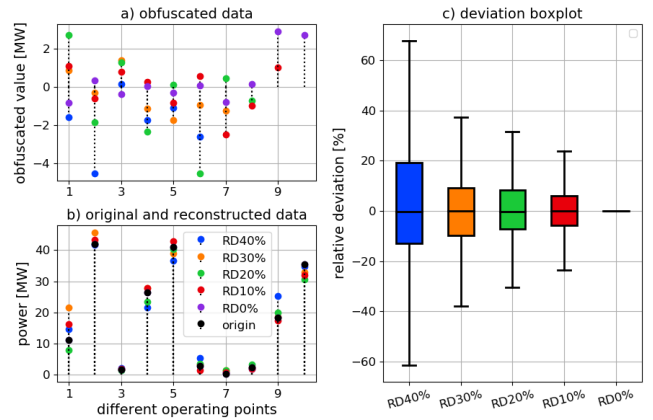


FIGURE 7. Comparison of obfuscated data with different reduction degrees: a) obfuscated values; b) reconstructed values; c) statistic deviation between reconstructed data set and original data set.

and RD20% (e.g., purple and green points in Fig. 7 a)) are used to investigate different RD effects on the accuracy of the equivalent.

3) ANN MODELS AND TRAINING RESULTS

After successful training with the real grid data set of the first eight months, the ANN-models in Table 8 and Table 9 are evaluated via the estimation for the last four months, see Fig. 6. It can be seen (left) that with the use of obfuscated data and increasing DR, the estimation accuracy decreases. However, the improvements brought by RNN offset this deviation to some extent, i.e., the accuracy of *FOCE0* with obfuscated data (light orange) is very close to the best (light blue). Lower RD is more likely to be chosen for better accuracy.

Plot b) in Fig. 8 shows the accuracy changes over one day. The power exchange between the DSO and the TSO between 8 am (time step 32) and 3 pm (time step 60) is estimated with a WAPE of around 1-2%. The reason is that the operating points of the observed grid assets typically vary more significantly and regularly from 8 am to 3 pm. Correspondingly, the training data set for this time window covers almost all relevant grid conditions, enabling the ANN to learn the relations correctly. The accuracy is worse at other time steps due to the lack of a “meaningful” training data set and the “inaccurate” estimation for the values close to zero. The TSO-DSO exchange from September to December ranges from 34.4 MW to -180 MW. Values close to zero cause high deviations, e.g., it is estimated to feed -0.02 MW to one boundary bus, but the actual feed-in is 0.03 MW, yielding an

TABLE 8. ANN-models for vertical equivalent.

ID	Input size	Description
FCE	166	FNN based on the original Critical grid data set, with measurements Error
RCE	166	RNN based on the original Critical grid data set, with measurements Error
FOCE0	166	FNN based on the Obfuscated Critical grid data set with RD of 0%, with measurements Error
ROCE0	166	RNN based on the Obfuscated Critical grid data set with RD of 0%, with measurements Error
FOCE20	166	FNN based on the Obfuscated Critical grid data set with RD of 20%, with measurements Error
ROCE20	166	RNN based on the Obfuscated Critical grid data set with RD of 20%, with measurements Error

TABLE 9. Hyperparameters for the ANN-models in Table 8.

Models	Hyperparameters
FCE and (RCE)	Layer No./ Layer size/ Activation function
	1 / 380 / LeakyReLU (LSTM)
	2 / 580 / LeakyReLU (LSTM)
	3 / 6 / linear
FOCE0 and (ROCE0)	Layer No./ Layer size/ Activation function
	1 / 390 / Tahn
Obfuscator + Estimator	2 / 166 / linear
	+
	1 / 540 / PReLU (LSTM)
	2 / 6 / linear
FOCE20 and (ROCE20)	Layer No./ Layer size/ Activation function
	1 / 580 / ReLU
	2 / 136 / linear
Obfuscator + Estimator	+
	1 / 60 / ReLU (LSTM)
	2 / 6 / linear

estimation error of 250% for that time step, but has negligible impacts on the grid operation in this scenario, cf. Fig. 9. Without consideration of the absolute values smaller than 1 MW, the accuracy of the ANN is significantly increased, see Fig. 8 c).

4) INTERACTION EVALUATION

A time-series simulation with the real grid data set of the last four months is carried out, considering measurement errors. Only the ANN-models with recurrent architecture (RCE, ROCE0, ROCE20) are considered. Note that the ANN-obfuscators are trained and constructed at the DSO and the TSO, respectively, cf. Fig. 3, and they are only valid for the grid data of the DSO or the TSO, respectively. In ROCE0 and ROCE20, obfuscated grid data of the DSO and the TSO are always used. The DSO-TSO interactions are estimated by different equivalent models and realized as equivalent loads attached at the boundary buses of the TSO.

Fig. 9 shows the maximum deviations of bus voltage, line loading, and line loss in the TSO grid for 99% of the simulated time window. The resulting grid state deviations for RCE and ROCE0 are very small, i.e., bus voltage deviations up to about 0.005%, line loading deviations up to 2%, and line loss deviations up to 0.05%. The ANN-model ROCE20 estimates the DSO-TSO interaction worse due to the compressed data set with RD of 20%, which causes larger deviations (green). The REIrep-model (pink) exhibits the smallest deviations

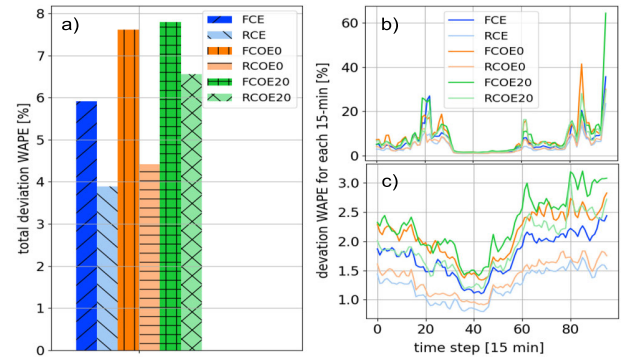


FIGURE 8. a) Estimation accuracy for the whole time series; b) Estimation accuracy each 15 min over one day; c) Estimation accuracy for each 15 min over one day without values smaller than 1 MW.

but still exhibits some deviations caused by the measurement errors. However, the shortcomings of the REIrep-model are distinct: as mentioned in Section IV-B-3), the DSO has to provide the recalculated REI equivalent model in every time step, requiring a complete grid model and probably a high computational burden. With realistically limited data sharing between the DSO and the TSO, e.g., only sharing PQ-operating points modification, the corresponding errors of the REI-model (purple), which is not recalculated in every time step, increase dramatically, i.e., to more than 1% voltage deviations, more than 20% line loading deviations, and more than 3% line loss deviations. These significant errors, in this scenario, are mainly caused by the time-varying tap changer positions that are only considered by the recalculation of an REI model based on the current and complete grid data of the DSO.

In contrast, under the same condition of limited data exchange, the ANN-models estimate the DSO-TSO interactions much more accurately. Although their effects on the TSO operation are still worse than that of the REIrep-model, the estimation errors are small, e.g., the line loading errors are all around 1-2%. Despite using an additional ANN model for data obfuscation (Autoencoder), the computational burden of the ANN equivalents is still almost instantaneous and substantially lower than that of the REIrep-model, cf. Table 6. In practice, the ANN-obfuscator guarantees the grid data confidentiality and cybersecurity in the process of data exchange, with a slight effect on equivalent accuracy.

5) DISCUSSION AND CONTRIBUTION TO DATA EXCHANGE ISSUES

There are different levels in which the exchange of grid data/model is necessary. Apart from the TSO-DSO exchange in this scenario, exchanges can be realized between TSOs and DSOs. The involved grid operators aim to perform common studies using shared data. The Common Information Model standard has helped to standardize the exchanges of grid data/model [43]. However, there are many methods for constructing this data or models. Different models describing the same networks and equipment may use different hierarchies, classes, attributes, associations, and nomenclature. These differences in modeling make it challenging to ensure

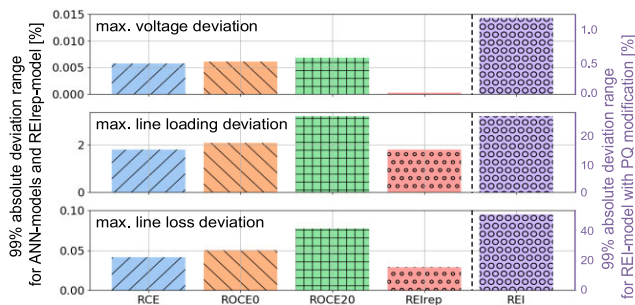


FIGURE 9. Estimation accuracy for ANN-estimators based on different data sets for 99% of the simulated time frame.

consistency of the network model between each usage [43]. The proposed ANN-based approach “skips” the constructing and power flow calculation processes, and the interoperability inside business processes is met through the ANN-estimation. Meanwhile, the use of our approach is more flexible (configurable inputs and outputs) and confidential (data obfuscation).

In another context, to build a (regional or pan-European) common grid model, the European TSOs share their information with regional security coordinators (RSCs). Based on the resulting broader overview of electricity, RSCs provide TSOs different services such as calculation of cross-border capacities, outage planning coordination for relevant transmission facilities *et al.* With the help of the proposed approach and based on the common grid model, RSCs could make their services efficient, e.g., in the future, TSOs can obtain/provide required information with lower efforts (sharing of limited grid data).

V. CONCLUSION AND OUTLOOK

In this paper, an innovative machine-learning-based grid equivalent approach is proposed for reducing an external grid area and equivalently representing its effect on an internal grid area accurately, adaptively, and confidentially. It is flexibly extendable, and the belonging components and tools can be applied independently, e.g., the ANN-obfuscator can benefit other studies related to confidential data.

The practical implementation of the proposed approach is expected to be relatively simple and efficient. There are no additional requirements for hardware. Meanwhile, the ANN-obfuscator makes the data exchange (exchange of an array or a JSON file consisting of obfuscated grid data) convenient and confidential. The proposed equivalent grid can be adapted to different power system studies according to the types of considered data, e.g., an ANN trained with switching status can be applied for congestion management. Nowadays, power systems are changing rapidly and significantly. In order to obtain an accurate equivalent model, the timeliness of the training data is essential. It is necessary to regularly retrain the ANN with the new grid data set to cope with the continuous changes of power systems. Thus, further improvement is needed to update the training data and retrain the ANN model automatically.

As an outlook, the compatibility of an ANN-based equivalent grid with more sophisticated grid operation studies, e.g., optimal power flow, is worth studying. Furthermore, this paper concentrated on the technical use of static equivalents for grid operations. It could also be applied and modified for grid planning and market-oriented studies.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support received from the University of Kassel, funds for open access publications, Kassel, Germany. The authors are fully responsible for the content of this publication.

REFERENCES

- [1] European Commission. (2015). *Energy Union Package—Achieving the 10% Electricity Interconnection Target Making Europe’s Electricity Grid Fit for 2020*. [Online]. Available: https://ec.europa.eu/energy/sites/ener/files/publication/FOR%20WEB%20energy%20union%20interconnections_EN-1.pdf
- [2] S. M. Ashraf, B. Rathore, and S. Chakrabarti, “Performance analysis of static network reduction methods commonly used in power systems,” in *Proc. 18th Nat. Power Syst. Conf. (NPSC)*, Dec. 2014, pp. 1–6, doi: 10.1109/NPSC.2014.7103837.
- [3] D. Shi and D. J. Tylavsky, “An improved bus aggregation technique for generating network equivalents,” in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2012, pp. 1–8, doi: 10.1109/PESGM.2012.6344668.
- [4] A. Papaemmanouil and G. Andersson, “On the reduction of large power system models for power market simulations,” in *Proc. 17th Power Syst. Comput. Conf. (PSCC)*, Aug. 2011, pp. 1308–1313.
- [5] J. Siermanns, M. Hotz, W. Utschick, D. Hewes, and R. Witzmann, “Feature- and structure-preserving network reduction for large-scale transmission grids,” in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–6, doi: 10.1109/PTC.2019.8810704.
- [6] M. Gavrilas, O. Ivanov, and G. Gavrilas, “REI equivalent design for electric power systems with genetic algorithms,” *WSEAS Trans. Circuits Syst.*, vol. 7, no. 10, pp. 911–921, Oct. 2008.
- [7] CEDEC, E.DSO, ENTSO-E, EURELECTRIC, and GEODE. (2019). *TSO-DSO Report: An Integrated Approach to Active System Management*. [Online]. Available: https://www.edsoforsmartgrids.eu/wp-content/uploads/2019/04/TSO-DSO_ASM_2019_190304.pdf
- [8] A. Zegers and H. Brunner. (Sep. 2014). *TSO-DSO Interaction: An Overview of Current Interaction Between Transmission and Distribution System Operators and an Assessment of Their Cooperation in Smart Grids*. [Online]. Available: https://www.iea-isgan.org/wp-content/uploads/2014/02/ISGAN_DiscussionPaper_TSODSOInteractionOverview_2014.pdf
- [9] Y. Phulpin, M. Begovic, M. Petit, J.-B. Heyberger, and D. Ernst, “Evaluation of network equivalents for voltage optimization in multi-area power systems,” *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 729–743, May 2009, doi: 10.1109/TPWRS.2009.2016534.
- [10] T. E. D. Liacco, S. C. Savulescu, and K. A. Ramarao, “An on-line topological equivalent of a power system,” *IEEE Trans. Power App. Syst.*, vol. PAS-97, no. 5, pp. 1550–1563, Sep. 1978, doi: 10.1109/TPAS.1978.354647.
- [11] J. B. Ward, “Equivalent circuits for power-flow studies,” *Trans. Amer. Inst. Electr. Eng.*, vol. 68, no. 1, pp. 373–382, Jul. 1949, doi: 10.1109/T-AIEE.1949.5059947.
- [12] S. Deckmann, A. Pizzolante, A. Monticelli, B. Stott, and O. Alsac, “Studies on power system load flow equivalencing,” *IEEE Trans. Power App. Syst.*, vol. PAS-99, no. 6, pp. 2301–2310, Nov. 1980, doi: 10.1109/TPAS.1980.319798.
- [13] F. C. Aschmoneit and J. F. Verstege, “An external system equivalent for on-line steady-state generator outage simulation,” *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 3, pp. 770–779, May 1979, doi: 10.1109/TPAS.1979.319289.
- [14] W. L. Snyder, “Load flow equivalent circuits—An overview,” in *Proc. IEEE PES Winter Meeting*, New York, NY, USA, Jan. 1972, pp. 436–444.
- [15] Z. Liu, N. Bornhorst, S. Wende-von Berg, and M. Braun, “A grid equivalent based on artificial neural networks in power systems with high penetration of distributed generation with reactive power control,” in *Proc. Conf. Sustain. Energy Supply Energy Storage Syst.*, Hamburg, Germany, Sep. 2020, pp. 1–7.

- [16] P. Dimeo, *Nodal Analysis of Power Systems*, Abacus Press, Kent, England, USA: Harwood Academic, 1975.
- [17] X. Cheng and T. J. Overbye, "PTDF-based power system equivalents," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1868–1876, Nov. 2005, doi: [10.1109/TPWRS.2005.857013](https://doi.org/10.1109/TPWRS.2005.857013).
- [18] H. Oh, "A new network reduction methodology for power system planning studies," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 677–684, May 2010, doi: [10.1109/TPWRS.2009.2036183](https://doi.org/10.1109/TPWRS.2009.2036183).
- [19] A. Papaemmanouil and G. Andersson, "On the reduction of large power system models for power market simulations," in *Proc. 17th Power Syst. Comput. Conf. (PSCC)*, San Diego, CA, USA, Aug. 2011, pp. 1–8.
- [20] H. K. Singh and S. C. Srivastava, "A reduced network representation suitable for fast nodal price calculations in electricity markets," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Jun. 2005, pp. 2070–2077.
- [21] G. Irisarri, A. M. Sasson, and J. F. Dopazo, "Real-time external system equivalent for on-line contingency analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 6, pp. 2153–2171, Nov. 1979, doi: [10.1109/TPAS.1979.319415](https://doi.org/10.1109/TPAS.1979.319415).
- [22] A. Marot, I. Guyon, B. Donnot, G. Dulac-Arnold, P. Panciatici, M. Awad, A. O'Sullivan, A. Kelly, and Z. Hampel-Arias. (Jun. 2020). *L2RPN: Learning to Run a Power Network in a Sustainable World NeurIPS2020 Challenge Design*. [Online]. Available: <https://www.public.asu.edu/~yweng2/Tutorial5/pdf/111.pdf>
- [23] J.-H. Menke, N. Bornhorst, and M. Braun, "Distribution system monitoring for smart power grids with distributed generation using artificial neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 113, pp. 472–480, Dec. 2019.
- [24] Z. Liu, S. Wende-von Berg, G. Banerjee, N. Bornhorst, T. Kerber, A. Maurus, and M. Braun, "Adaptives statisches Netzäquivalent mit künstlichen neuronalen Netzen," in *Proc. Symp. Energieinnovation*, Graz, Austria, Feb. 2020, pp. 1–11.
- [25] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, "Hyperband: A novel bandit-based approach to hyperparameter optimization," 2016, *arXiv:1603.06560*.
- [26] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent., (ICLR)*, San Diego, CA, USA, May 2015, pp. –15.
- [27] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 47–57, Mar. 2017, doi: [10.1109/TCI.2016.2644865](https://doi.org/10.1109/TCI.2016.2644865).
- [28] *Pytorch Documentation*. Accessed: May 20, 2021. [Online]. Available: <https://pytorch.org/docs/stable/generated/torch.nn.MSELoss.html>
- [29] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. De-Vito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *Proc. NIPS Autodiff Workshop*, Long Beach, CA, USA, Dec. 2017, pp. 1–4.
- [30] L. Thurner, A. Scheidler, F. Schafer, J.-H. Menke, J. Dollichon, F. Meier, S. Meinecke, and M. Braun, "Pandapower—An open-source Python tool for convenient modeling, analysis, and optimization of electric power systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6510–6521, Nov. 2018.
- [31] *DigSILENT GmbH, Power Factory*. (2020). Accessed: May 20, 2021. [Online]. Available: <http://www.digsilent.de>
- [32] V. Stankovic, *Introduction to Machine Learning in Image Processing*. Glasgow, U.K.: Univ. Strathclyde, 2017.
- [33] J. Si, A. Barto, W. Powell, and D. Wunsch, "Reinforcement learning and its relationship to supervised learning," in *Handbook of Learning and Approximate Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2004, pp. 45–63, doi: [10.1109/9780470544785.ch2](https://doi.org/10.1109/9780470544785.ch2).
- [34] R. Sathya and A. Abraham, "Comparison of supervised and unsupervised learning algorithms for pattern classification," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 2, pp. 34–38, Feb. 2013.
- [35] G. Chen, "A gentle tutorial of recurrent neural network with error back-propagation," 2016, *arXiv:1610.02583*.
- [36] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D, Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306.
- [37] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [38] Z. Wang, S. Wende-von Berg, and M. Braun, "Fast parallel Newton-Raphson power flow solver for large number of system calculations with CPU and GPU," *Sustain. Energy, Grids Netw.*, vol. 27, Sep. 2021, Art. no. 100483.
- [39] M. Yousefi-Azar, V. Varadharajan, L. Hamey, and U. Tupakula, "Autoencoder-based feature learning for cyber security applications," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 3854–3861.
- [40] A. Dertat. (2017). *Applied Deep Learning—Part 3: Autoencoders*. [Online]. Available: <https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798>
- [41] S. Meinecke, D. Sarajlić, S. R. Drauz, A. Klettke, L.-P. Lauen, C. Rehtanz, A. Moser, and M. Braun, "SimBench—A benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis," *Energies*, vol. 13, no. 12, p. 3290, Jun. 2020.
- [42] C. Chen, J. Twycross, and J. M. Garibaldi, "A new accuracy measure based on bounded relative error for time series forecasting," *PLoS ONE*, vol. 12, no. 3, pp. 1–23, 2017, doi: [10.1371/journal.pone.0174202](https://doi.org/10.1371/journal.pone.0174202).
- [43] ENTSO.E. (May 2014). *Common Grid Model Exchange Standard (CGMES)*. [Online]. Available: <https://www.entsoe.eu/digital/cim/cim-for-grid-models-exchange/>
- [44] D. Bogen, G. Latisko, and K. Dziegielewski, "The journey to the centralized CIM based network data model management at ONCOR," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2012, pp. 1–5, doi: [10.1109/PESGM.2012.6345087](https://doi.org/10.1109/PESGM.2012.6345087).



ZHENG LIU received the B.Sc. and M.Sc. degrees in electrical engineering from the Technical University of Ilmenau, Ilmenau, Germany, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Department of Energy Management and Power System Operation, University of Kassel, Kassel, Germany. His research interests include grid equivalents, reactive power control, and machine learning in general.



JAN-HENDRIK MENKE received the bachelor's and master's degrees in electrical engineering and information technology from TU Dortmund University, Dortmund, Germany, and the Ph.D. degree in electrical engineering from the University of Kassel, Kassel, Germany, in 2019. He is currently working for the transmission system operator Amprion with the System Development Department.



NILS BORNHORST received the Diploma and Ph.D. degrees in electrical engineering from the Darmstadt University of Technology, Darmstadt, Germany, in 2009 and 2014, respectively. He is currently the Group Leader with the Department of Energy Management and Power System Operation, University of Kassel. His research interests include optimization and machine learning methods for the operation of power grids with a high share of renewables.



MARTIN BRAUN (Senior Member, IEEE) received the Diploma degree in electrical engineering and in technically oriented business administration from the University of Stuttgart, Stuttgart, Germany, in 2005, and the Ph.D. degree in engineering from the University of Kassel, Kassel, Germany, in 2008. He is currently a Professor with the Department of Energy Management and Power System Operation, University of Kassel, and the Director of the Grid Planning and Grid Operation

Division, Fraunhofer Institute for Energy Economics and Energy System Technology, Kassel.

• • •