

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

A Robust Preprocessing and Topology-Aware Feature Engineering Framework for Power System State Estimation under Sparse and Outlier-Affected Measurements

Said Ćosić¹ and István Vokony¹

¹Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary

Corresponding author: Said Cosic (e-mail: saidcosic@edu.bme.hu).

ABSTRACT This paper presents a comprehensive, AI-based state estimation framework designed to address the complexities introduced by reduced inertia, fast-changing topologies and erratic measurements. The proposed framework introduces a layered preprocessing and feature-engineering workflow that consists of data preprocessing techniques and advanced feature engineering that embeds spatial and temporal context to the integration of topological awareness and machine learning-based estimation. A custom-designed artificial neural network is trained with tailored optimization strategies and regularization mechanisms to enhance convergence, deliver high estimation accuracy and strong resilience to bad or corrupted data. Additionally, robust solver techniques are integrated to further improve estimation reliability under challenging conditions. The proposed framework achieves consistently high accuracy across networks of varying size, redundancy, and renewable penetration, outperforming the conventional SE, especially in low-observability grids. Robustness testing confirms superior performance under bad data and topology errors, with success rates up to 25% higher than the conventional method, though gross measurement errors remain challenging. Sensitivity analysis shows that the algorithmic solver, temporal and spatial features, and robust preprocessing are the strongest contributors to both accuracy and robustness, with their influence increasing as measurement sparsity intensifies.

INDEX TERMS Artificial intelligence, artificial neural network, feature engineering, state estimation.

I. INTRODUCTION

Accurate and robust power system monitoring has been recognized as one of the essential prerequisites for normal power system operation [1]. However, challenges introduced as a part of global energy transition, such as bidirectionality of smart grids, have created additional complexities in maintaining situational awareness and preserving continuity of electricity supply [2]. Large-scale proliferation of distributed energy resources, advanced control mechanisms, more complex communication networks and high proliferation of power electronics, have introduced more frequent and new forms of stability issues, posing challenges for traditional monitoring mechanisms.

State estimation, one of the most significant aspects of power system monitoring, has traditionally relied on phasor measurement units (PMU), physical models and linearized system equations. While effective in steady-state conditions, these systems often lack the responsiveness and accuracy

needed for dynamic real-time assessment, especially in the presence of noisy data, complex interactions, and rapidly evolving contingencies [3]. In addition to this, conventional methods fail to take into account new potential forms of instability, such as topology-induced instability as a consequence of dynamic reconfigurations due to fast switching, microgrid operation, as well as cybersecurity-based threats [4].

These factors underscore the urgent need for smarter, more flexible, accurate and reliable state estimation frameworks to achieve optimal system performance and system stability. For these reasons, this research aims to design and develop a state estimation framework capable of enhanced stability monitoring that successfully handles the introduced complexities and nonlinearities and achieves a considerably higher level of system controllability and automation.

II. STATE ESTIMATION IN POWER SYSTEMS

State estimation is a cornerstone of grid monitoring and control in modern power systems. The following sections offer a brief overview of fundamental concepts related to state estimation and its functions, as well as an elaboration of key issues and challenges addressed by this paper.

A. DEFINITION AND MATHEMATICAL FORMULATION

One of the most significant objectives in everyday power system operation is maintaining the normal state, which heavily relies on the ability of SCADA systems to continuously monitor the system through the acquisition of various measurements, their processing and determining the system state [5]. This process of inferring the values of system state variables using a limited number of measured data at certain locations is referred to as state estimation (SE) [6]. State estimation is a vital component of energy management systems (EMS) and essentially represents a numerical transformation that infers state variables from a subset of various measured data [7].

Mathematically, state estimation is most commonly formulated as an overdetermined system of nonlinear equations and solved as an optimization problem with a quadratic objective function, also referred to as Weighted Least Square (WLS), with equality and inequality constraints [8].

Consider the nonlinear measurement model:

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} h_1(x_1, x_2, \dots, x_n) \\ h_2(x_1, x_2, \dots, x_n) \\ \vdots \\ h_m(x_1, x_2, \dots, x_n) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = h(\mathbf{x}) + \mathbf{e} \quad (1)$$

where $\mathbf{z}$ is the measurement vector, $h_i(\mathbf{x})$ is the nonlinear scalar function relating measurement i to the state vector, $\mathbf{e}$ is the measurement error vector, which is assumed to have zero mean and variance, σ_i^2 , and $[x_1, x_2, \dots, x_n] = \mathbf{x}^T$ is the system state vector, $\mathbf{x}$. There are m measurements and n state variables, $n < m$.

If this problem is formulated in terms of measurement residuals: $r_i = z_i - h(z_i)$, the problem becomes:

$$\begin{aligned} \text{minimize } J(\mathbf{x}) &= \frac{1}{2} \sum_{i=1}^m \frac{r_i^2}{\sigma_i^2} \\ \text{subject to } g_j(\mathbf{x}) &= 0, \quad j = 1, \dots, n_g \\ c_j(\mathbf{x}) &\leq 0, \quad j = 1, \dots, n_c \end{aligned} \quad (2)$$

where g_j and c_j are functions representing power flow quantities.

Whereas the most common state variables are nodal voltage magnitude and angle, transformer turns ratio magnitude and angle, as well as complex active and reactive power flow, the most common measurement variables are bus and line active and reactive power flows, bus voltage magnitudes and line currents [9].

B. STATE ESTIMATION FUNCTIONS

State estimation process includes several interconnected modules, each of which has a specific function and application. Their interactions are shown in Fig. 1.

In general, the process starts with gathering sensor input data, where measurements are aggregated by data concentrators [10], and is followed by topology identification, in which the network configuration is verified and updated based on the status of switching elements, such as circuit breakers and disconnectors, to mitigate potential topological inconsistencies [11]. After this, observability analysis determines the feasibility of uniquely estimating the state variables given the current measurement set and identifies any unobservable regions as a result of insufficient measurement redundancy [12].

If insufficient measurement redundancy is detected, pseudo-measurement generation module is activated to create the missing measurements from historical or statistical models and improve network observability [13]. Upon confirming observability, the SE solver computes the optimal set of system states by solving a constrained optimization problem that incorporates the network model and the measurement data. The process continues with bad data processing, which aims to detect, isolate, and if necessary, replace erroneous measurements [14]. The final stage involves the integration of SE results into downstream processes of power system monitoring and control, such as stability assessment, contingency analysis, model validation and optimal power flow.

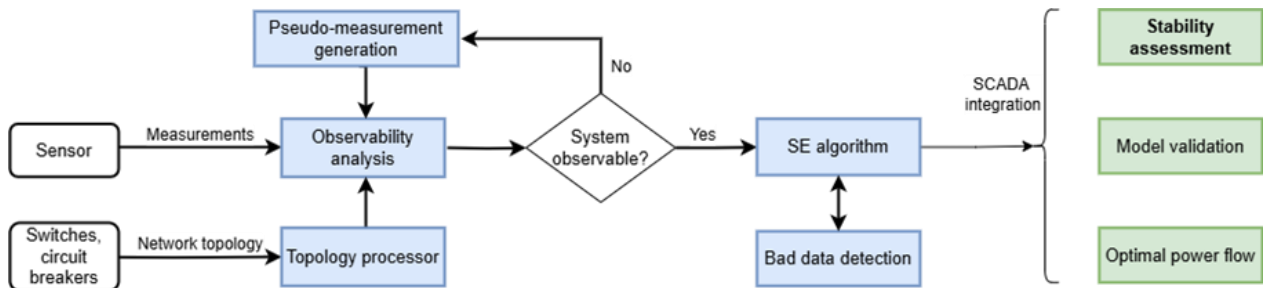


FIGURE 1. Typical State Estimation workflow. Process steps are denoted by blue boxes, and some possible applications are represented in green boxes.

1) TOPOLOGY PROCESSING

Network Topology Processing (NTP) is a module that aims to identify energized, de-energized, and grounded electrical islands and is usually performed before state estimation and other related functions [15].

In conventional NTP [16], raw analog measurements are initially converted into standardized engineering units and system conditions, such as operational limits, logical switching constraints and rate-of-change thresholds are verified. Following this, bus section processing is performed to determine electrical connectivity by identifying topological islands in the form of bus section groups. The final step involves aggregating the physical bus sections into the bus/branch model. A variant of NTP that is capable of selectively updating only the parts of the network affected by changes in switching device statuses in the so-called tracking mode is presented in [17]. Generalized NTP approach [18] further improves the conventional NTP by integrating the topology identification of extended islands into the estimation process and explicitly representing uncertain or erroneous switching device statuses.

Recently proposed methods build upon these premises by formulating the topology verification as a maximum likelihood estimation problem [19] in a statistical learning framework for verifying single-phase distribution grid topologies using non-synchronized voltage data from smart meters. Other methods achieve high topology processing accuracy by modeling the grid as a graph and identifying subgraphs likely to contain topology errors in a Nonlinear Least Absolute Value (NLAV) framework [20] or by employing a probabilistic method such as Recursive Bayesian Approach (RBA) [21].

2) OBSERVABILITY ANALYSIS

Observability analysis determines if a unique SE solution exists, i.e., if the state variables can be inferred from the available measurement set [5], and identifies parts of the network which are unobservable due to bad or missing data, as well as observable islands on which SE can be applied [22]. The methods of verifying system observability are generally classified into numerical, topological and hybrid approaches.

Whereas the most widely adopted conventional techniques of branch variables, nodal variables and topological observability are based on the graph theory of spanning trees and work on the principle of formulating every grid topology as a system of loop equations, novel research approaches are based on probabilistic methods (such as Gaussian Mixture Model [23], multivariate complex Gaussian modeling [24]) and machine learning (e.g. Artificial Neural Network [25]).

3) PSEUDO-MEASUREMENT GENERATION

In case input measurement redundancy is inadequate and the system is unobservable, the input set is artificially expanded by introducing pseudo-measurements, which are usually generated according to the network's historical data by various methods. Two main categories of these methods are

distinguished in literature: probabilistic and statistical approaches, and learning-based approaches. Whereas probabilistic methods involve Gaussian Mixture Models (GMM) of weighted, finite sum of Gaussian probability density functions [26], Kernel Density Estimation (KDE) [27], Expectation Maximization (EM) [28], and Schweppe-Huber Generalized Method (SHGM) based on projection statistics [29], learning-based approaches rely on the concepts of Parallel Distributed Processing (PDP) networks [30], a game-theoretic expansion of Relevance Vector Machines (RVM) [31], and Gradient Boosting Trees (GBT) [32].

4) SE SOLVER

The state estimation solver aims to find an optimal solution for the system states according to the network model constraints and given measurements. The optimal solution comprises of a vector of complex bus voltages, according to which estimates are calculated for other network values, such as generator outputs, loads, line flows, transformer taps, etc.

Conventionally, the solver includes a linear optimization problem and contains an algorithm that minimizes an objective function in an iterative way. One of the most common approaches to determine the system states is the statistical method of maximum likelihood estimation (MLE), which uses the logical assumption that measurement data follows a specific probability distribution and calculates parameters of the probability density function so that the function has maximum value.

Solutions to the optimization problem can be obtained using either Newton's method [33] or the iterative re-weighted least squares (IRWLS) approach [34]. Whereas Newton's method solves a set of nonlinear equations derived from the Lagrangian via Taylor expansion, iteratively updating the gradient matrix, IRWLS introduces a penalty factor to reduce the influence of outliers and repeatedly applies WLS with updated weights. In addition to these, several recent methods SE solver have been proposed that utilize Semi-Definite Programming (SDP) [35] and linear state estimation [36].

5) BAD DATA DETECTION

Bad data detection usually consists of two major steps: bad data detection, a procedure of deciding whether the measurement data set includes any bad data, and bad data identification, which involves determining which particular measurements contain bad data.

Conventional algorithms rely on the assumption that measurement residuals follow a particular probability density function and perform statistical hypothesis testing on the data that exceeds the threshold ϵ [37]. The most frequently used techniques in this group are Chi-squares test [38], normalized residuals approach [39], and L2 norm testing [40].

Novel bad data detection approaches have utilized linear WLS frameworks with equality constraints [41], largest normalized residual test of PMU data [42], and advanced statistical analyses based on penalized semidefinite programming conic relaxation [43].

Recently, it has been shown that the model-based detection algorithms described above are susceptible to attack vectors created by adversaries with prior knowledge of the system [44]. To address these vulnerabilities, a wide range of data-driven methods has been proposed, the most prominent of which seem to be supervised learning techniques [45], Support Vector Machine (SVM) [46], and k-means clustering (KMC) [47].

C. KEY ISSUES AND CHALLENGES

Recent advancements in static and dynamic state estimation have revealed several persistent challenges that limit the effectiveness of current methodologies. Some of the key concerns are numerical stability, convergence, robustness against bad data, accuracy of pseudo-measurements, as well as design of efficient real-time SE framework.

Numerical stability and convergence are particularly significant when integrating measurements with differing accuracy and scales from diverse sources like RTUs, PMUs, and DERs. The biggest challenges in this domain are optimality and observability in systems with marginal measurement redundancy [48]. Despite improvements in bad data handling, new potential threats, such as data corrupted by cyberattacks, necessitate more resilient SE frameworks. Furthermore, employing new methods to generate pseudo-measurements to enhance SE accuracy is a significant concern, especially when considering the cost associated with the installation of advanced metering equipment, especially in the distribution systems [7]. From a performance standpoint, the growing system complexity necessitates scalable and distributed SE architecture capable of real-time application, in addition to a more standardized approach that promotes easier implementation to future adopters.

This paper addresses the above issues and proposes solutions of the above-mentioned challenges by utilizing data-driven techniques based on artificial intelligence (AI).

III. AI-BASED DYNAMIC STATE ESTIMATION FRAMEWORK

The main aim of the proposed framework is to exploit deficiencies of conventionally used SE aspects, with a particular focus on enhancing pseudo-measurement generation and state estimator capabilities by utilizing AI techniques and feature engineering.

A. BASIC AI PRINCIPLES

AI refers to the theory and development of computational systems capable of performing tasks that typically require human intelligence, such as reasoning, learning, decision-making, perception, and language understanding [49]. The most basic principle of AI is its ability to imitate or replicate cognitive functions of humans, enabling computers to analyze data, draw conclusions, and act upon them with minimal human intervention. According to [49], the ultimate goal of AI

is to create systems that can autonomously perceive their environment, process information, and act adaptively to achieve specific objectives.

AI systems are generally structured as intelligent agents, which perceive their environment through sensors that collect data, interpret information by using internal models based on a variety of learning algorithms, make decisions through reasoning capabilities, and act upon the environment through actuators to achieve specific goals [50].

AI encompasses a wide set of methods aimed at enabling machines to simulate intelligent behavior. The most common classification of AI approaches is based on learning paradigm, whereby there are three major types of AI frameworks: supervised learning – where models are trained on labeled data where the correct output is provided for each input with the aim of minimizing error, unsupervised learning – in which models find pattern or structures in unlabeled data, such as clustering or dimensionality reduction, and Reinforcement Learning (RL) methods, in which decision-making is defined as a Markov Decision Process models where models learn by interacting with an environment and receiving feedback in the form of rewards or penalties with the aim to maximize cumulative reward over time [51].

One of the major AI techniques is based on Artificial Neural Network (ANN), which is a feedforward, layered architecture comprised of multiple layers of interconnected neurons that apply nonlinear activation functions, such as ReLU, sigmoid, or tanh, to their weighted combinations of inputs. Their typical structure consists of input, hidden, and output layers of neurons such that neurons of one layer connect only to neurons of the immediately preceding and immediately following layers. Each weight represents a neuron's relative importance. The training process is based on minimizing a loss function (cumulative prediction error of neural network) via backpropagation and gradient-based optimizers (e.g., SGD, Adam) [52].

B. PROPOSED FRAMEWORK

Fig. 2 illustrates a comprehensive flowchart of proposed methodologies used for pseudo-measurement generation (PMG), designed to enhance power system state estimation through data-driven methods and feature optimization. The process begins with the collection of input data, which includes static and dynamic system information and grid topology data. From this foundation, case studies and specific scenarios are selected to define the operating conditions under which the model will be developed and tested. The next phase involves network modeling and simulation, after which a quasi-dynamic simulation is performed using a time-series of load, generator and RES profiles, to produce input and output measurements that are processed (by cleaning, normalization, and structuring). As the most crucial, feature engineering and topological considerations are applied to enhance model learning.

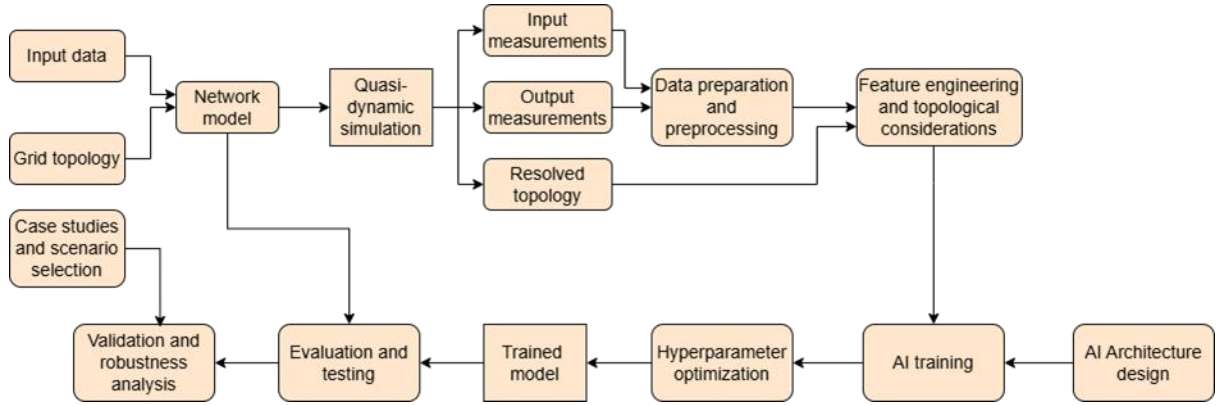


FIGURE 2. Flowchart of applied research methodology.

Then the prepared data is used to train the AI model and hyperparameter optimization is used to ensure that the model performs efficiently and generalizes well. Finally, the trained model is subjected to evaluation and testing to assess its accuracy, and sensitivity analysis is conducted to assess impacts of the proposed enhancements. Additionally, validation and robustness analysis are conducted to examine the model's performance under different operating conditions and topological configurations.

C. INPUT DATA PREPARATION AND PREPROCESSING

The input grid data consisting of three sample systems was provided by the SimBench dataset [53]: EHV-HV grid (Grid A), HV-MV grid (Grid B), MV-LV grid (Grid C). Table I displays basic network characteristics for each of the grids in terms of number of elements. Whereas Grid A is a replica of the German 380- and 220-kV grid, grid B represents a suburban HV and MV open ring system, and grid C emulates a typical rural MV-LV topology.

For each of the grids, in addition to the base case scenarios, denoted as A0, B0 and C0, respectively, there are two additional cases, denoted as A1, A2, B1, B2, C1 and C2, respectively. The cases differ in the allocations of renewable generation capacity per fuel type, such that the capacity progressively increases, as illustrated in Fig. 3 for Grid C.

The load profiles are modelled as a time series of 35136 steps (yearly data in 15-minutes resolution) with normalized active and reactive power data for each load type. Each load has an assigned type, such as rural, urban, commercial etc. and each load type possesses its own load curve in each grid.

TABLE I
BASIC CHARACTERISTICS OF TEST GRIDS

Element Type	Grid A	Grid B	Grid C
No. of lines	1057	204	5391
No. of loads	523	156	5373
No. of terminals	713	173	5483
No. of conventional generators	345	0	0
No. of RES generators	422	192	581
No. of substations	32	19	1
No. of transformers	218	8	92

Similarly, all generator types (including non-RES, RES and storage) are modelled with normalized active power profiles.

The data preparation process for the training procedure involved interfacing DIgSILENT PowerFactory software [54] with Python 3.8 and creating time characteristics for all input data profiles of load, generator, RES and storage, and running a quasi-dynamic analysis on the three test grids for the time-series simulation of 35136 time intervals.

The quasi-dynamic simulation yielded the output consisting of the bus voltage angles and magnitudes at each bus, an example of which is illustrated in Fig. 4, for grid C, allowing the initial training input and output data to be taken from the PowerFactory variables: active power flow, reactive power flow, voltage magnitude, and voltage angle.

The data was additionally processed by imputation of historical averages of the input measurements for weekday-hour pairs, since it was observed that the input measurements have a high weekly seasonality. Thus, an additional input was created considering the average value of all P and Q node measurements of the same hour and the same weekday, so that e.g. all Thursday 10:00h measurements of active power at a single bus have the same average value.

Furthermore, in order to make the data more resistant to outliers and random distributions, the data vectors were normalized by using the robust scaling approach:

$$X_{scaled} = \frac{X - \text{median}(X)}{IQR(X)} \quad (3)$$

where IQR is the difference between 75th percentile (Q3) and 25th percentile (Q1). After the scaling, all data points outside the criterion: $|X_{scaled}| > 2$ were cleared as obvious outliers.

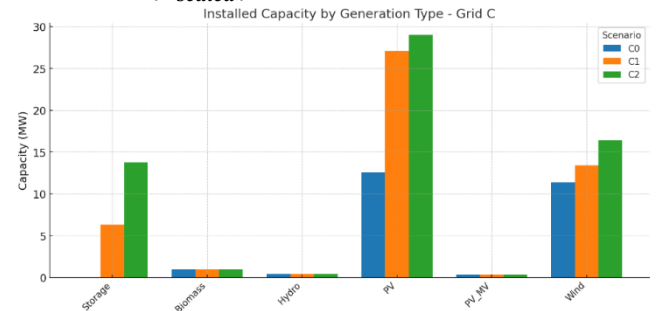


FIGURE 3. Installed generation capacity [MW] per fuel type in Grid C

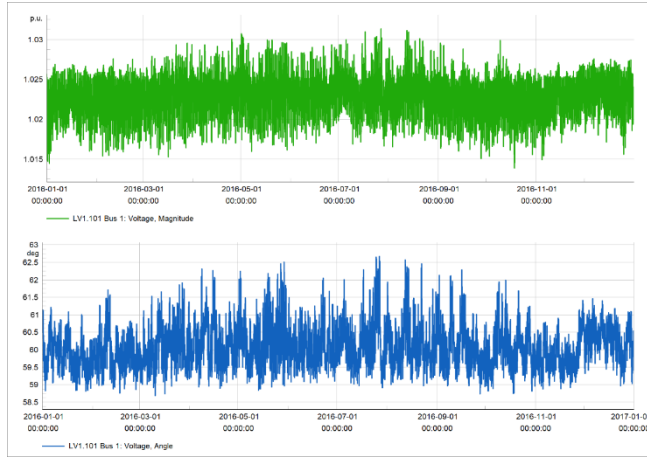


FIGURE 4. Yearly voltage magnitude and angle at bus LV1.101.

An additional enhancement of input data (which was used in robustness testing and not in the original training) was done by simulating sparse conditions. This included artificially generating a Boolean mask of the same shape as the data matrix with the rates of [0.2, 0.4, 0.6] which produces a matrix of uniform random numbers in [0,1) and randomly scatters them, replacing the original measurements with zeros to emulate sensor outages or communication dropouts.

D. FEATURE ENGINEERING AND TOPOLOGICAL CONSIDERATIONS

To help the model uncover underlying patterns and relationships and further improve accuracy, four methods of feature engineering were proposed, some of which highly rely on input or output topology results:

1) Integration of temporal features

The input data timestamps were expanded to offer a better insight into cyclic time patterns of the data. The following temporal indicators were proposed: hour-of-day, day-of-week,

and season. They are represented as integer types with the ranges of [1-24], [1-7], and [1-4], respectively.

2) Augmentation of spatial context

Since each element in the input dataset contains information about the subnet parameter that corresponds to the grid to which the element belongs, the output set was enhanced by adding extra parameters such as the average of voltage magnitude of all buses belonging to the same subnet (PowerFactory object zone).

3) Reduction of output measurement outliers

To further reduce eventual measurement outliers, a method is applied which sorts all voltage measurements in descending order and computes their standard deviations. Following this, measurements from the bottom or top of the list are removed and the new standard deviation is computed. If the new standard deviation is found to be less than half of the previously calculated one, the measurement is substituted by the mean value of the measurement set that doesn't include the outlier. This procedure is repeated until removing a measurement no longer results in a reduction of the standard deviation by more than 50%.

4) Inclusion of topological considerations

This extra feature involves the topology solver's output data, particularly the node energization information obtained from PowerFactory's time-series load flow results (variable named e:ciEnergized). Consequently, a binary indicator that indicates if a specific node is energized or not is created and appended to the dataset, enabling a much better capturing of network connectivity dynamics by removing the need for phantom predictions, as well as enhancing overall robustness, generalization and potential failure mode identification by distinguishing between normal load variations and fault-induced islanding events.

A summary of all proposed data enhancement indicators is shown in Table II.

TABLE II
SUMMARY OF FEATURE ENGINEERING ENHANCEMENT INDICATORS

Indicator	Type	Stage	Group	Source	Target variables
V_prev	float	Preprocessing	Historical average imputation	input measurements	P, Q for training
X_scaled	integer	Preprocessing	Robust scaling	input & output measurements	P, Q, V for training
Z_bad	boolean matrix	Preprocessing	Sparse condition simulation	input measurements	P, Q for testing
Hour	integer	Feature engineering	Temporal features	timestamps	P, Q for training
Day	integer	Feature engineering	Temporal features	timestamps	P, Q for training
Season	integer	Feature engineering	Temporal features	timestamps	P, Q for training
V_zone_avg	float	Feature engineering	Spatial context	input topology, output measurements	V for training
V_bad	n/a	Feature engineering	Output outliers reduction	output measurements	V for training
Energized	boolean	Feature engineering	Topological consideration	output topology	P, Q, V for training

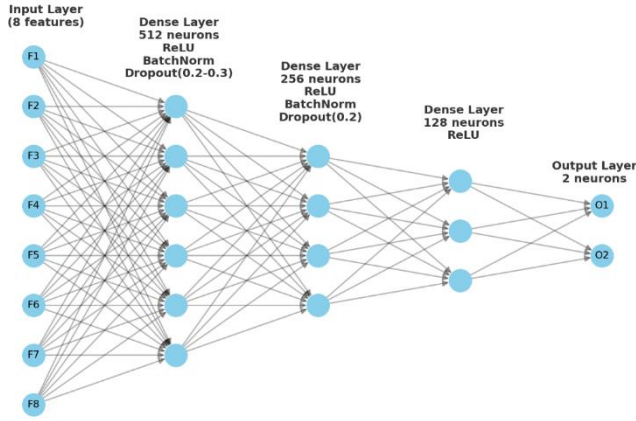


FIGURE 5. Diagram of the optimal ANN structure. It consists of an input layer, 3 hidden layers and an output layer in an 8-512-256-128-2 configuration.

E. AI DESIGN AND TRAINING

an ANN-based feed-forward perceptron architecture is proposed. As previously established, ANNs are computational networks which attempt to simulate the decision process in networks of nerve cells (neurons) of the biological central nervous system [55]. A perceptron, which is a commonly used model for a neuron, can be defined by the following input-output relation:

$$z = \sum_i w_i x_i \quad (4)$$

$$y = f_A(z) \quad (5)$$

where z is the node (summation) output of linearly combined weighted (w_i) inputs, x_i , used as an argument of activation function, f_A , which is typically linear, unipolar, binary or sigmoid. Typically, an ANN is composed of several perceptrons arranged in parallel to form a layer, with one or more such layers connected in sequence. In this structure, known as a multi-layer perceptron, the output from every neuron in one layer is transmitted to each neuron in the subsequent layer.

The proposed multi-layered perceptron structure allows for one ANN network to estimate the bus voltage magnitudes, and a second one to estimate the voltage angles. The input layer size is equal to the product of the number of input measurements and the added features, and the numbers of critical buses. The output layer size is equal to the number of bus voltage magnitudes or angles, respectively. Initially, three hidden dense layers are proposed for each of the test grids, and the Rectified Linear Unit (ReLU) as their activation function, to promote sparse, efficient representations. Moreover, the initial setup proposes batch normalization in the first and second dense layers to normalize the inputs to zero mean and unit variance, accelerating convergence and stabilizing training, as well as a dropout rate of 20% to prevent overfitting and encourage redundancy. Finally, the output layer provides simultaneous regression outputs: voltage magnitude and angle per bus.

The AI training procedure, which utilizes a Python 3.8 environment's scikit-learn library for machine learning [56],

starts with determination of the dataset split ratio. In this research, the split of 75-25% is applied, which means that 75% of the data is used for training and 25% for validation. The training data comprises a number of training samples and features for each sample, where the total dataset size is determined by the product of sample and feature number. However, due to samples being time-series results, the augmented data having the size of $(\text{time steps} \times (\text{no. of features} \cdot \text{no. of buses}))$ needed to be transformed to avoid an input layer of too large size and possible overfitting errors. For this, the data science technique of melting, i.e., converting a wide-format dataset into a long-format one by unpivoting columns into rows, was utilized. As a result, the training dataset size of $(4 \cdot 24 \cdot 366 \cdot n_b) \cdot 10$ was reached, with n_b defined as the number of buses in a subset of buses taken from the test grid. The features set consists of an array of 4 measurements (2 input, 2 output) and 6 additional features explained in the previous sections. The Adam optimization algorithm is employed to adjust the ANN's weights, initially using a learning rate of 0.001, with the objective of minimizing the mean squared error (MSE) between the network's output and the target values. In the process, the early stopping option is used, meaning that the validation error is monitored to stop training early if no improvement is reached in several consecutive epochs, i.e., complete cycles through the entire training dataset.

F. HYPERPARAMETER OPTIMIZATION

Unlike model parameters (weights and biases), hyperparameters are set prior to training and critically influence convergence speed, generalization ability and computational efficiency. Hyperparameters such as the number of layers, the number of neurons per layer, learning rate, batch size, dropout rate, the number of training epochs and the amount of training data are fine-tuned by calculating them using several combinations and recording the success rate. After the initial training was conducted on a large, diverse grid (Case A), the fine-tuning assessments were performed on the MV (Case B) grid with less sample data. The ideal number of layers is chosen based on heuristics and the input dataset considerations. To balance model capacity with generalization, three hidden layers are chosen. For the ideal number of neurons per layer, estimations were made for the combinations represented in Table III. The chosen configuration, shown in Fig. 5, is the one with the lowest average validation p.u. mean square error (MSE).

TABLE III
VALIDATION MSE FOR DIFFERENT ANN CONFIGURATIONS.

ANN configuration (number of neurons in hidden layers)	Average validation MSE
32-16-8	0.089
64-32-16	0.08
128-64-32	0.072
256-128-64	0.068
512-256-128	0.075

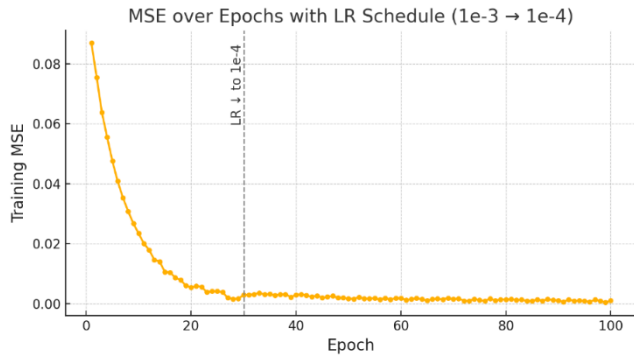


FIGURE 6. MSE progression with hybrid LR schedule approach.

Since the initial learning rate of 0.001 has proven to be slightly too large, and the minimum validation MSE has occurred around 0.0001, but using it may lead to slower convergence, a hybrid schedule-based learning rate has been proposed in which the learning rate changes from the initial 0.001 to 0.0001 after the 30th epoch, as can be seen in Fig. 6.

For the training dataset, the most optimal maximum epoch number has proven to be between 50 and 150, and therefore, the middle value of 100 is chosen. Moreover, the batch size of 256 and the dropout rate of 20% are chosen as a rule of thumb, without further detailed analysis.

IV. RESULTS

As training and optimization have been conducted on diverse networks with different generation and load patterns, as well as topological structures, but on a grid that is inherently observable, the capability of the proposed approach to produce accurate pseudo-measurements needs to be tested by using a range of scenarios that feature low measurement redundancy and/or low observability. Moreover, robustness to bad data such as gross errors, missing measurements, and topology inaccuracies has been tested on different scenarios. Finally, sensitivity analysis has been conducted to assess the contributions of the proposed feature enhancements.

A. EVALUATION AND TESTING

The accuracy of the proposed approach was tested on the grids described in Table I. Table IV shows the number of installed measurements in each of the defined cases, indicating that grids B1, B2, C0, C1 and C2, have the lowest measurement redundancy in addition to having the high penetration of renewables.

TABLE IV
MEASUREMENT CONFIGURATION THROUGHOUT THE CASE STUDIES

Meas. type	Case								
	A0	A1	A2	B0	B1	B2	C0	C1	C2
<i>i</i>	952	986	1058	65	2	2	2	2	2
<i>p</i>	1533	1291	1633	321	36	38	14	14	16
<i>q</i>	1533	1291	1633	321	36	38	14	14	16
<i>v</i>	708	597	599	284	54	55	7	7	9
Total	4726	4165	4923	991	128	133	7	7	9

TABLE V
COMPARISON OF SE ACCURACY

Case	Proposed framework			Conventional SE		
	MAE (pu)	RMSE (pu)	R ²	MAE (pu)	RMSE (pu)	R ²
A0	0.0021	0.003	0.974	0.002	0.0029	0.961
A1	0.0019	0.0028	0.989	0.002	0.0031	0.976
A2	0.0021	0.0031	0.953	0.0022	0.003	0.956
B0	0.0026	0.0035	0.989	0.0033	0.0042	0.98
B1	0.0012	0.0021	0.956	0.0015	0.0025	0.946
B2	0.0015	0.0024	0.968	0.0019	0.0029	0.958
C0	0.0022	0.003	0.988	0.0028	0.0036	0.979
C1	0.0024	0.0029	0.977	0.003	0.0035	0.967
C2	0.0016	0.0023	0.973	0.002	0.0028	0.963

To compare the proposed SE method with the conventional SE with respect to accuracy, another WLS-based SE is performed on the initial measurement and pseudo-measurement datasets. The following error metrics are used to assess the performance of the proposed method and benchmark with the conventional SE:

- Mean Absolute Error (MAE) - the average of the absolute differences between predicted and actual values,
- Root Mean Squared Error (RMSE) - the square root of the average squared differences between predicted and actual values,
- Coefficient of Determination (R² Score) - the proportion of variance in the actual data that is predictable from the model.

The estimation results are shown in Table V, across all error metrics and consider average error values for both output measurement vectors, as tested by 9 case studies of the 3 test networks. The results demonstrate that the ANN-based SE maintains consistently low error levels and high determination coefficient, even as grid size, measurement density, generation and load profiles and storage capacity vary. In the large EHV-HV network (A-cases), MAE hovers around 0.0021 p.u. for baseline and maximum storage case (A0 and A2, respectively), dipping to 0.0019 when storage is doubled. The R² peaking to 0.989 in A1 and falling to 0.953 in A2 suggests diminishing returns once storage exceeds certain point despite slightly higher measurement counts. In the MV-LV grid (B-cases) with 100% RES penetration, baseline case (B0) displays the highest MAE and RMSE, but an excellent R², and the addition of storage in B1 and B2 seems to improve the MAE by roughly half and reduce the RMSE by around 30%, followed by a small decrease in R². Finally, in the MV-LV network (C-cases) MAE and RMSE errors remain modest, even with very sparse measurements, while R² stays above 0.97. Moreover, the results indicate that the ANN-based SE performs comparably well in grids with high measurement redundancy (Cases A), whereas for medium-redundancy suburban networks (Cases B) and low-redundancy rural networks (Cases C), its accuracy outperforms the conventional method across all metrics.

TABLE VI
ROBUSTNESS TESTING SCENARIOS

Case ID	Type	Description
<i>M1</i>	Gross Measurement Error	+50% voltage offset at HV1 Bus 1
<i>M2</i>	Gross Measurement Error	+50% voltage offset at MV4.101 Bus 76
<i>M3</i>	Gross Measurement Error	+50% P, Q offset at MV4.101 Bus 15
<i>O1</i>	Measurement Outage	Voltage measurement of 0 p.u. at HV1 Bus 1
<i>O2</i>	Measurement Outage	Voltage measurement of 0 p.u. at MV4.101 Bus 76
<i>O3</i>	Measurement Outage	P, Q measurement of 0 at HV1 Bus 15
<i>T1</i>	Topology Error	Switch status of HV1 Switch 1 inverted
<i>T2</i>	Topology Error	Switch status of MV1 Switch 1 and 2 inverted
<i>T3</i>	Topology Error	R, X of line MV4.101 loop_line 1 90 % of actual value

The accuracy of the conventional SE deteriorates more due to increased sensitivity to measurement sparsity and noise, with relative differences of up to 25% compared to the proposed method.

B. ROBUSTNESS ANALYSIS

The objective of robustness analysis is to assess the influence of different operating conditions on estimation accuracy, considering variations in measurement quantity and type, as well as the method's robustness to bad data such as gross errors, missing measurements, and topology inaccuracies. For that purpose, different test scenarios are developed, as described in Table VI, and tested on the B2 scenario (MV-HV grid, full renewable generation, storage, DER, with extremely low measurement redundancy of around 15%). The cases include an increase of bus voltage measurement at random buses, emulating a sensor outage by setting the voltage measurement as 0 p.u., inverting the binary status of switches (open/close), and decreasing line resistance and reactance parameters by 10%. The success condition of each scenario is to achieve a voltage magnitude prediction error of less than 1% or a voltage angle prediction error of less than 5%. The success rate is measured as the ratio of all scenarios where the tested method satisfies the success condition to the maximum number of successful tests (i.e., for each of the 9 robustness scenarios there are 18 tests = 9 case studies $\times$ 2 variables). Finally, the proposed framework is compared against the conventional SE across the robustness metrics (M1–M3, O1–O3, T1–T3) in terms of success rate, and the results are presented in Fig. 7.

From the figure, it may be observed that the proposed framework dominates in every category, often by a large margin, especially in T1, T2 and T3 (topological errors). This can be attributed to the feature engineering and preprocessing done on the ANN input data.

Proposed vs. Conventional SE Framework Performance

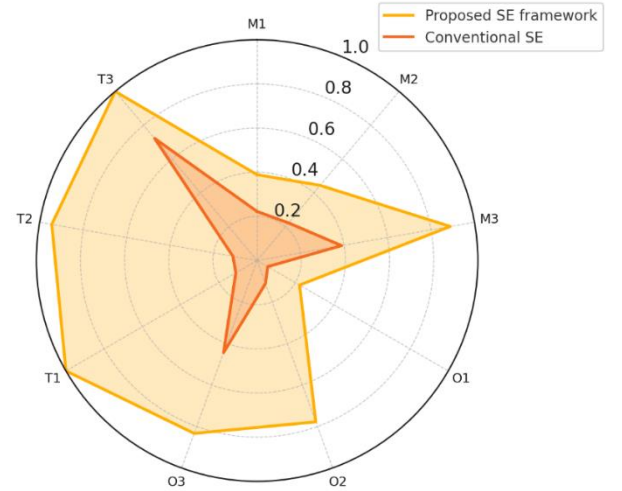


FIGURE 7. Comparison of robustness of the SE methods.

Moreover, O1 shows a modest gain, suggesting a similar performance of both methods assuming the strict 1% criterion, in the case of HV measurement outage. However, the other two cases of measurement outages (voltage outage at an MV bus and a P, Q outage) prove that the proposed approach significantly improves robustness. Furthermore, for gross measurement errors, the proposed framework fails to offer a significantly large success rate, which can be explained with the strict success rate and the requirement of convergence.

C. SENSITIVITY ANALYSIS

To assess the relative contribution of each framework component to the overall performance gains sensitivity analysis is conducted as a series of ablation studies that systematically remove key elements of the proposed framework, such as the feature engineering enhancements from Table II in such a way that only one enhancement is added at a time. Consequently, the ANN is retrained on a smaller input set and tested on the test cases C1, C2 and C3. Additionally, a case in which the SE solver is replaced by the Schweppe–Huber solver and IRWLS algorithm is applied on the ANN-generated pseudo-measurements is tested. Table VII displays the results in terms of average relative difference in RMSE (compared to the results in Table V) and the relative difference in the success rate of the robustness testing scenarios (comparison with Table VI) for each of the cases.

From the table, it can be seen that all the proposed components display performance enhancements and generally have a bigger influence as the measurement sparsity increases. Furthermore, the largest single contributor to the RMSE improvement is the SE solver, followed by temporal and spatial features. Whereas some features, such as robust scaling and historical average imputation show a particularly stronger effect under noisy conditions, other features, such as topological considerations show a moderate but consistent impact.

TABLE VII
SENSITIVITY ANALYSIS RESULTS OF THE PROPOSED FEATURE ENGINEERING ENHANCEMENT INDICATORS

Indicator	Stage	Group	Relative RMSE difference [%]			Relative SR difference [%]								
			C1	C2	C3	M1	M2	M3	O1	O2	O3	T1	T2	T3
V_{prev}	Preprocessing	Historical average imputation	3.2	4.1	7.9	6.1	5.8	6.5	14.3	15.1	13.8	3.5	3.9	4.1
X_{scaled}	Preprocessing	Robust scaling	8.4	12.3	15.1	18.4	17.3	19.1	10.2	10.7	9.9	5.8	6.1	6.4
Z_{bad}	Preprocessing	Sparse condition simulation	2.1	4.2	6	4.2	4.6	4.4	9.5	10.2	9.8	3.1	3.4	3.6
$Hour$	Feature engineering	Temporal features												
Day	Feature engineering	Temporal features	10.5	10.4	17.8	12.8	12.2	13.5	9.1	9.6	8.9	7.3	7.8	8.1
$Season$	Feature engineering	Temporal features												
V_{zone_avg}	Feature engineering	Spatial context	5.3	11.1	12.2	8.6	9.2	8.9	7.4	7.7	7.1	13.4	14.1	14.7
V_{bad}	Feature engineering	Output outliers reduction	4.1	7.3	3.0	10.1	10.7	10.3	7.6	8.1	7.3	6	6.4	6.7
$Energized$	Feature engineering	Topological consideration	4.2	3.8	9.1	6.5	6.8	6.4	5.8	6.1	5.7	17.5	18.3	17.9
SH and $IRWLS$	SE solver	Algorithm	12.1	16.8	20.3	20.7	19.9	21.3	15.9	16.4	15.1	12.1	12.7	13.4

Moreover, the biggest factor affecting robustness is the algorithm, as it provides the highest universal SR improvement across all cases. It is also noteworthy that different feature enhancements have shown stronger influence (biggest SR gains) on different categories of the robustness analysis, such as robust scaling and algorithm choice on the gross errors (M1-M3), historical imputation and sparse condition simulation on the outages (O1-O3). Similarly, topological considerations and spatial context addition provide the highest boosts of the topology errors.

V. CONCLUSION

This paper presents a cohesive end-to-end framework that combines robust data preprocessing, advanced feature design and ANN-learning-based pseudo-measurement generation. To counteract measurement gaps and erratic outliers, a layered preprocessing and feature engineering workflow is developed. By combining statistics-based imputations, IQR normalization, iterative cleaning of extreme values, and adversarial injection of synthetic data dropouts, the framework successfully reconstructs a reliable, high-dimensional input space. The addition of spatial averages and topology-derived connectivity flags further embeds structural knowledge, resulting in a robust observability platform that operates effectively even when the majority of measurements are unavailable or compromised.

Therefore, as demonstrated on test grids, the proposed methodology achieves high estimation accuracy, effectively capturing the system states across various operating conditions and significantly outperforming the conventional SE methodology in terms of robustness against sudden data fluctuations and measurement errors.

ACKNOWLEDGMENT

István Vokony acknowledges the support of the Bolyai János Research Scholarship of the Hungarian Academy of Sciences (BO/50/24).

REFERENCES

- [1] J. Zhao *et al.*, "Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, Jul. 2019, doi: 10.1109/TPWRS.2019.2894769.
- [2] 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, 2019, doi: 10.1016/j.ijepes.2019.05.057.
- [3] J. Qi, K. Sun, J. Wang, and H. Liu, "Dynamic State Estimation for Multi-Machine Power System by Unscented Kalman Filter With Enhanced Numerical Stability," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1184–1196, 2018, doi: 10.1109/TSG.2016.2580584.
- [4] B. Tan, J. Zhao, V. Terzija, and Y. Zhang, "Decentralized data-driven estimation of generator rotor speed and inertia constant based on adaptive unscented Kalman filter," *Int. J. Electr. Power Energy Syst.*, vol. 137, p. 107853, 2022, doi: https://doi.org/10.1016/j.ijepes.2021.107853.
- [5] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. 2004.
- [6] A. Monticelli, "State Estimation in Electric Power Systems. A Generalized Approach," *Proceedings of the IEEE*. 1999.
- [7] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *IEEE Transactions on Smart Grid*. 2019, doi: 10.1109/TSG.2018.2870600.
- [8] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, 2000, doi: 10.1109/5.824004.
- [9] J. Zhao *et al.*, "Roles of dynamic state estimation in power system modeling, monitoring and operation," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2462–2472, May 2021, doi: 10.1109/TPWRS.2020.3028047.
- [10] O. Darmis and G. Korres, "A Survey on Hybrid SCADA/WAMS State Estimation Methodologies in Electric Power Transmission Systems," *Energies*, vol. 16, no. 2, 2023, doi: 10.3390/en16020618.

- [11] M. Asprou, S. Chakrabarti, and E. Kyriakides, "A Two-Stage State Estimator for Dynamic Monitoring of Power Systems," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1767–1776, 2017, doi: 10.1109/JSYST.2014.2375951.
- [12] M. Glavic and T. Van Cutsem, "Reconstructing and tracking network state from a limited number of synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1921–1929, 2013, doi: 10.1109/TPWRS.2012.2231439.
- [13] E. Manitsas, R. Singh, B. Pal, and G. Strbac, "Modelling of pseudo-measurements for distribution system state estimation," in *CIREN Seminar 2008: SmartGrids for Distribution*, 2008, pp. 1–4, doi: 10.1049/ic:20080447.
- [14] A. S. Musleh, G. Chen, and Z. Y. Dong, "A Survey on the Detection Algorithms for False Data Injection Attacks in Smart Grids," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2218–2234, 2020, doi: 10.1109/TSG.2019.2949998.
- [15] M. H. Vuong, S. Lefebvre, and X. D. Do, "Detection and estimation of topology and parameter errors from real-time measurements," in *IEEE Power Engineering Society Summer Meeting*, 2002, vol. 3, pp. 1565–1569 vol.3, doi: 10.1109/PSS.2002.1043653.
- [16] K. G.N. and K. P.J., "Reduced model for numerical observability analysis in generalised state estimation," *IEE Proc. - Gener. Transm. Distrib.*, vol. 152, no. 1, pp. 99–108, Jan. 2005, doi: 10.1049/ip-gtd:20041059.
- [17] M. Prais and A. Bose, "A topology processor that tracks network modifications," *IEEE Trans. Power Syst.*, vol. 3, no. 3, pp. 992–998, 1988, doi: 10.1109/59.14552.
- [18] N. Singh and H. Glavitsch, "Detection and identification of topological errors in online power system analysis," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 324–331, 1991, doi: 10.1109/59.131079.
- [19] G. Cavarero, V. Kekatos, and S. Veeramachaneni, "Voltage Analytics for Power Distribution Network Topology Verification," *IEEE Trans. Smart Grid*, 2019, doi: 10.1109/TSG.2017.2758600.
- [20] S. Park, R. Mohammadi-Ghazi, and J. Lavaei, "Topology error detection and robust state estimation using nonlinear least absolute value," 2019, doi: 10.23919/acc.2019.8814813.
- [21] B. Hayes, A. Escalera, and M. Prodanovic, "Event-triggered topology identification for state estimation in active distribution networks," 2016, doi: 10.1109/ISGTEurope.2016.7856295.
- [22] Y. H. Moon, Y. M. Park, and K. J. Lee, "Observable island identification for state estimation using incidence matrix," *Automatica*, vol. 24, no. 1, pp. 71–75, 1988, doi: [https://doi.org/10.1016/0005-1098\(88\)90008-8](https://doi.org/10.1016/0005-1098(88)90008-8).
- [23] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, 2010, doi: 10.1109/TPWRS.2009.2030271.
- [24] N. Iqteit, A. Basa Arsoy, and B. Çakır, "Load Profile-Based Power Loss Estimation for Distribution Networks," *Electrica*, vol. 18, no. 2, pp. 275–283, 2018.
- [25] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, 2012, doi: 10.1109/TPWRS.2012.2187804.
- [26] S. R., P. B.C., and J. R.A., "Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement," *IET Gener. Transm. Distrib.*, vol. 4, no. 1, pp. 50–59, Jan. 2010, doi: 10.1049/iet-gtd.2009.0167.
- [27] Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "A Probabilistic Data-Driven Method for Photovoltaic Pseudo-Measurement Generation in Distribution Systems," 2019, doi: 10.1109/PESGM40551.2019.8974026.
- [28] D. Teymouri, O. Sedehi, L. S. Katafygiotis, and C. Papadimitriou, "A Bayesian Expectation-Maximization (BEM) methodology for joint input-state estimation and virtual sensing of structures," *Mech. Syst. Signal Process.*, vol. 169, p. 108602, 2022, doi: <https://doi.org/10.1016/j.ymssp.2021.108602>.
- [29] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics [of power systems]," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, 1996, doi: 10.1109/59.496203.
- [30] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1008–1016, 2013, doi: 10.1109/TPWRS.2012.2215927.
- [31] K. Dehghanpour, Y. Yuan, Z. Wang, and F. Bu, "A Game-Theoretic Data-Driven Approach for Pseudo-Measurement Generation in Distribution System State Estimation," *IEEE Trans. Smart Grid*, 2019, doi: 10.1109/TSG.2019.2893818.
- [32] Z. Cao, Y. Wang, C. C. Chu, and R. Gadh, "Robust pseudo-measurement modeling for three-phase distribution systems state estimation," *Electr. Power Syst. Res.*, 2020, doi: 10.1016/j.epsr.2019.106138.
- [33] M. Cosovic and D. Vukobratovic, "Distributed Gauss-Newton Method for State Estimation Using Belief Propagation," *IEEE Trans. Power Syst.*, 2019, doi: 10.1109/TPWRS.2018.2866583.
- [34] R. C. Pires, A. Simoes Costa, and L. Mili, "Iteratively reweighted least-squares state estimation through Givens Rotations," *IEEE Trans. Power Syst.*, vol. 14, no. 4, pp. 1499–1507, 1999, doi: 10.1109/59.801941.
- [35] T. Ratnarajah, Z.-Q. Luo, and K. M. Wong, "Semidefinite programming solutions to robust state estimation problem with model uncertainties," in *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No.98CH36171)*, 1998, vol. 1, pp. 275–276 vol.1, doi: 10.1109/CDC.1998.760683.
- [36] L. Zhang and K. Lai, "A novel complex linear state estimator for smart power distribution systems: methodology and implementation," *Int. J. Electr. Power Energy Syst.*, 2020, doi: 10.1016/j.ijepes.2020.106312.
- [37] J. Angulo-Paniagua and J. Quiros-Tortos, "Comparing Chi-square-based bad data detection algorithms for distribution system state estimation," 2020, doi: 10.1109/TDLA47668.2020.9326241.
- [38] M. Göl and A. Abur, "A modified Chi-Squares test for improved bad data detection," in *2015 IEEE Eindhoven PowerTech*, 2015, pp. 1–5, doi: 10.1109/PTC.2015.7232283.
- [39] A. S. Dobakhshari, V. Terzija, and S. Azizi, "Normalized Deleted Residual Test for Identifying Interacting Bad Data in Power System State Estimation," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 4006–4016, 2022, doi: 10.1109/TPWRS.2022.3144316.
- [40] J. G. Sreenath, A. Meghwani, S. Chakrabarti, K. Rajawat, and S. C. Srivastava, "A recursive state estimation approach to mitigate false data injection attacks in power systems," in *2017 IEEE Power & Energy Society General Meeting*, 2017, pp. 1–5, doi: 10.1109/PESGM.2017.8274070.
- [41] A. Jovicic and G. Hug, "Linear state estimation and bad data detection for power systems with RTU and PMU measurements," *IET Gener. Transm. Distrib.*, 2020, doi: 10.1049/iet-gtd.2020.0487.
- [42] H. Khazraj, F. Faria Da Silva, and C. L. Bak, "Comparison between conventional and post-processing PMU-based state estimation to deal with bad data," 2017, doi: 10.1109/IEEEIC.2017.7977482.
- [43] R. Madani, J. Lavaei, R. Baldick, and A. Atamtürk, "Power System State Estimation and Bad Data Detection by Means of Conic Relaxation," 2017, doi: 10.24251/hicss.2017.375.
- [44] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, Jun. 2011, doi: 10.1145/1952982.1952995.
- [45] S.-C. Yip, K. Wong, W.-P. Hew, M.-T. Gan, R. C.-W. Phan, and S.-W. Tan, "Detection of energy theft and defective smart meters in smart grids using linear regression," *Int. J. Electr. Power Energy Syst.*, vol. 91, pp. 230–240, 2017, doi: <https://doi.org/10.1016/j.ijepes.2017.04.005>.
- [46] H. Yinghui and L. Shiqi, "Detection of Bad Data and Cyber Attacks in State Estimation using Support Vector Machine," in *2022 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2022, pp. 111–115, doi: 10.1109/ISGTAsia54193.2022.10003475.
- [47] T. Zhang, H. Li, Y. Huang, M. Zhai, and F. Zeng, "Bad Data Detection and Identification of Hybrid AC/DC Power Systems with Voltage Source Converters Using Deep Belief Network and K-Means Clustering," *Elektron. ir Elektrotechnika*, vol. 25, no. 2 SE-, pp. 60–66, Apr. 2019, doi: 10.5755/j01.eie.25.2.23206.
- [48] L. Kamyabi, T. T. Lie, S. Madanian, and S. Marshall, "A Comprehensive Review of Hybrid State Estimation in Power Systems: Challenges, Opportunities and Prospects," *Energies*, vol.

- 17, no. 19, 2024, doi: 10.3390/en17194806.
- [49] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (4th Edition)*. Pearson, 2020.
- [50] J. Shuford and M. M. Islam, "Exploring the Latest Trends in Artificial Intelligence Technology: A Comprehensive Review," *J. Artif. Intell. Gen. Sci.* ISSN3006-4023, vol. 2, Feb. 2024, doi: 10.60087/jaigs.v2i1.p13.
- [51] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [52] U. Shahzad, "Artificial Neural Network for Transient Stability Assessment: A Review," *2024 29th Int. Conf. Autom. Comput.*, pp. 1–7, 2022.
- [53] S. Meinecke *et al.*, "SimBench-A benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis," *Energies*, vol. 13, no. 12, 2020, doi: 10.3390/en13123290.
- [54] DIgSILENT, "PowerFactory," <https://www.digsilent.de/en/powerfactory.html> (accessed May 01, 2025).
- [55] D. Graupe, *Principles of Artificial Neural Networks*. 2016.
- [56] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, 2011.



SAID ĆOSIĆ received the B.Sc. degree (magna cum laude) in electrical and electronics engineering from International University of Sarajevo (IUS), Sarajevo, Bosnia and Herzegovina, in 2018, and the M.Sc. degree (summa cum laude) in electrical engineering from Budapest University of Technology and Economics (BME), Budapest, Hungary, in 2020.

He has served as a Power System Modelling and Analysis Engineer as well as a Power System Data Exchange Consultant in Budapest, Hungary.

His research interests encompass power system stability, smart grids, renewable energy integration, and advanced system modelling.

Mr. Ćosić is currently a Ph.D. candidate at the Department of Electrical Power Engineering at Faculty of Electrical Engineering and Informatics, BME. He has been awarded a Stipendium Hungaricum M.Sc. and Ph.D. scholarship in recognition of his academic achievements.



ISTVÁN VOKONY (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree from the Budapest University of Technology and Economics (BME), in 2007 and 2012, respectively.

He has been with the Department of Electric Power Engineering at BME since 2007, where he is currently an Associate Professor. He has also held leadership positions at E.ON Business Services Hungary Ltd. and has been actively

involved in numerous national and international R&D projects, including several H2020 initiatives.

Dr. Vokony is currently the Head of the Department of Electrical Power Engineering at the Faculty of Electrical Engineering and Informatics, BME. He is also a former Officer of the AEE Hungary student chapter and an active member of the IEEE, the Hungarian Scientific Association of Energy (MEE), the Association of Energy Engineers (AEE), as well as a founding member of the World Energy Council's HYPE initiative. His research interests include system stability, renewable energy integration, energy efficiency, and smart grids.