State Estimator Degrading Observability for IEEE33 and IEEE37 Test Systems by using Graph Neural Networks

Vlad SANDULEAC
Dept. of Electrical Power Systems
University Politehnica of Bucharest
Bucharest, Romania
vlad.sanduleac@upb.ro

Remus TEODORESCU
Department of Energy Technology
Aalborg University
Aalborg, Denmark
ret@energy.aau.dk

Mihai SANDULEAC
Dept. of Electrical Power Systems
University Politehnica of Bucharest
Bucharest, Romania
mihai.sanduleac@gmail.com

***Abstract*—As distribution networks are undergoing a transformation towards Active distribution networks, and as Smart grid concepts are becoming increasingly present in such networks, distribution system state estimators (DSSE) are becoming also necessary to support monitoring and control. With the emergence of artificial intelligence and the introduction of Graph Neural Networks (GNN) for better capturing grid topologies and their properties, new frontiers are opened for applying this technique for real needs. The paper is assessing the Degrading Observability for a DSSE by analyzing the subject for two specific test grids – IEEE 33 and IEEE 37 when using GNN and shows robustness of results against inputs degrading observability.**

Keywords— State estimator, Graph Neural Network, Degrading observability, masking rate

I. INTRODUCTION

Observability of distribution grids is a main target in the active distribution grids, as they face various challenges due to high renewables penetration, consumption and storage needs, while there is also a need for efficient use of the grid and for increasing energy resilience. Smart metering and Smart grid efforts help incrementally to increase network observability, however complete physical observability is not economic, while partially implemented metering and PMUs programs may face temporary communication problems, defects and also malicious cyber-attacks. State estimators are meant to cover grid data which are not available or which have been distorted due to several reasons. The paper is aiming to analyze the robustness of a state estimator based on Graph Neural Networks with various degrees of "blindness" or unavailability of data, in terms of accuracy in estimating the missing data in distribution power grids. As the testing domain is very large, the paper considers a number of two IEEE distribution test grids for medium voltage level for assessing the observability degradation with various degradation levels.

The paper has the following structure: it gives the background and shows the relevance of studying DSSE with neural networks; presents the NN-based approach and the two specific test grids IEEE 33 and IEEE 37 to be used in the study; presents the results related to the state estimator degrading observability and its capability to provide DSSE output data for various masking rates of the input data.

II. BACKGROUND AND RELEVANCE

Distribution systems state estimation (DSSE) is an important tool to correct data acquired through SCADA systems and to increase observability by deducting values for unavailable data. DSSE is considered as an important step towards implementing Smart Grids paradigm [1]. Advanced DSSE are reviewed also in [2], where it is also shown the importance of considering also phasor measurement units for increasing accuracy. In [3] it is shown a combination of SCADA measurements and AMI pseudo measurement modeling method based on sliding time window to increase the total input data which improves the precision of the state estimator. The paper gives also a method to increase the input data by using AMI data in combination with real-time data from SCADA. Using SCADA, PMU and AMI multi-source data for DSSE is also reported in [4], with simulations on a IEEE-13 bus active distribution network. State estimation based on hybrid measurements of RTU, PMU and AMI is also presented in [5] and has been illustrated with numerical simulation using also an IEEE 13-bus system. Conventional Node Voltage based State Estimation (NVSE) and Branch Current based State Estimation (BCSE) have been analyzed and validated on a 33-bus distribution feeder [6], showing also the impact of the number of PMU measurements state estimation performance.

Neural network (NN) based State estimators have been also studied in various works. In [7], a convolutional neural network (CNN) based method is used for pseudo measurement modeling in order to achieve the system observability. The study [8] puts forward a state estimation method for distribution network based on improved residual neural network, showing better estimation accuracy and robustness than other methods. Unrolled Spatiotemporal Graph Convolutional Network has been proposed for DSSE in [9], with simulation results for 118-node and 1746-node distribution systems.

An open question is to determine the amount of data needed as observed inputs for DSSE versus the quality of the state estimator results. To address this issue, this paper studies the effect of degrading observability in state estimation and the manner in which the performance of DSSE is influenced by increasingly reduced observability. The model studied is a variation of Graph Neural Networks (GNN), while the degradation in observability has been simulated by masking input parameters at varying masking rates, using two specific test grids: IEEE-33 and IEEE-37.

III. NN-BASED APPROACH

The distribution grids are by their nature firstly described by their topology. This is also the reason why test grids are denominated by the number of nodes, which also implies a specific topology. With this in mind, the NN based state estimator implemented in this study has been chosen to be a GNN using graph attention. The attention mechanism is described in formula 1 [10]. The scope of this work is the implementation and training of a GNN State Estimator (GNN-

\

SE) in order to encapsulates the grid behavior well enough to be able to perform State Estimations for the part of the grid data which is marked as being unknown or suspicious.

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

A. Training Method

The models were trained in a masked modeling paradigm. In this paradigm, a proportion of the data is hidden probabilistically with a likelihood controlled by a masking rate M. Any given data point is hidden by replacing the latent space representing that data point inside the model with a learnable masked embedding. Thus, the model learns to approximate the masked reconstruction function $f_{\boldsymbol{\Theta}}$, as seen in the formula below, where $\mathcal{G}$ is the graph representation of the network, $\mathbf{m}$ is the observation mask, $\mathcal{O}$ are the observed values and $\mathcal{H}$ are the hidden values.

$$\hat{\mathbf{y}}_{\mathcal{H}} = f_{\boldsymbol{\Theta}}(\mathbf{y}_{\mathcal{O}}, \mathbf{m}, \mathcal{G}) \quad (2)$$

The learning rate was controlled by a cosine annealing with warm restarts scheduler to avoid local minima. The models were trained for 600 epochs, with checkpoints being saved every 50 epochs and at the best validation loss.

In this experiment, the masking rates M was varied from 50% to up to 95% to test the resilience of the proposed model to adverse training and inference conditions. The 95% masking rate was chosen as a maximum duress state.

B. Tracking the observability accuracy

For the purposes of this study, the parameters tracked were chosen to be U, U_{angle}, P_{bus}, P_{line}, Q_{bus}, Q_{line}, I_{bus}, I_{line}. Line variables represent powers and currents relating to lines between buses, while bus variables represent injected (or extracted) powers and currents relating to elements connected to the given bus.

Range-normalized mean absolute error (NMAE) in percent, calculated relative to the largest absolute value for a given parameter, was used alongside range-normalized variance σ for evaluation.

$$NMAE = 100\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{\max_j(|y_j|)}\right| \quad (3)$$

$$\sigma = \frac{100}{\max_j|y_j|}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - \bar{e})^2} \quad (4)$$

$$e_i = y_i - \hat{y}_i \quad (5)$$

These values are calculated in each training iteration (epoch) and represent main indicators of observability accuracy during the training and testing phase.

IV. USECASES IEEE33 AND IEEE37

The tests have been made on two different test grids: IEEE 33 and IEEE 37. In addition to the total number of nodes which briefly describe the topology of these IEEE test grids, three more properties have been also considered:

- The degree of ramification of each grid N_{Ramif}, which means the number of grid nodes where ramification occurs (at least 3 lines connected to the node mean that it is a node with ramifications).

- The minimal depth of branches $MinDepth_B$, which is given by the smallest number of branches between two ramification nodes or between the ramification and the end of a feeder.

- The maximal depth of branches $MaxDepth_B$, which is given by the highest number of branches between two ramifications or between the ramification and the end of a feeder.

A. The IEEE 33 test grid

The initial IEEE 33 test grid topology is presented in [11] and has been adapted for 20 kV voltage level, while the slack bus has been considered at the bottom of the HV/MV transformer, which is not anymore considered in the scheme used in this study. Similar adaptations have been made in other works, such as [12], where additional branches may be also introduced. The topology used for the current study is presented in Fig. 1.

This test grid has only three ramification nodes: 2, 3 and 6, thus $N_{Ramif} = 3$. The maximum depth of branches is $MaxDepth_B = 12$, corresponding to the feeder starting in node 6 and ending in node 18, which allows studies on a network with long feeders. The minimum depth of branches is $MinDepth_B = 3$, which is obtained between nodes 3 and 6 and also between nodes 3 and 25.

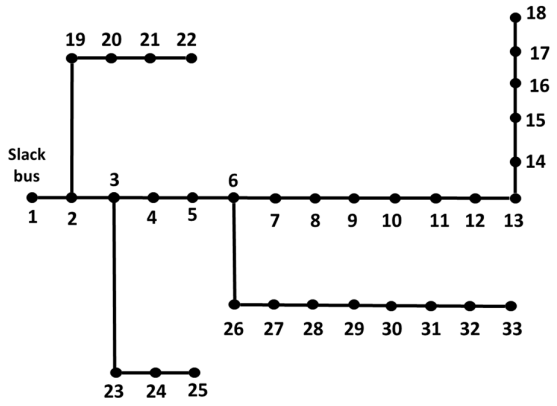


Fig. 1. The IEEE-33 distribution network, as used in this paper.

B. The IEEE 37 test grid

The grid is defined in [13], described in the context of other IEEE distribution test feeders in [14] and mentioned in several works ([15] and others). Grid topology and parameters can be found also at [16]. The topology is presented in Figure 2, while in this work, it has been adapted for 20 kV voltage level, specifically to European distribution grids.

This test grid is different from the previous one, having a higher number of ramification nodes $N_{Ramif} = 12$, which are the nodes 2, 3, 4, 5, 7, 8, 9, 10, 11, 16, 27 and 35.

The maximum depth of total branches $MaxDepth_B = 3$, given by the branch between nodes 27, 30, 31 and 11, which is much smaller than for IEEE 33. Moreover, the minimum depth of total branches $MinDepth_B = 1$, e.g. between nodes 2 and 3, but also in many other situations.

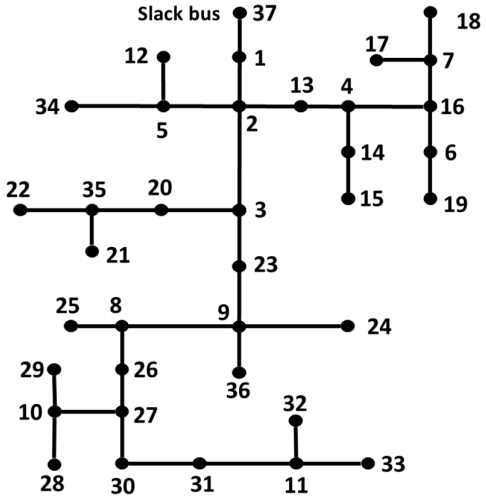


Fig. 2. The IEEE-37 distribution network.

In this manner, we are able to study a different pattern of the grid, while maintaining a similar complexity in terms of node number as compared to the first test grid (37 instead of 33).

C. *Test grids parameters comparison*

Based on the three parameters introduced before, a comparison between the two test grids is given in Table 1.

TABLE I. TEST GRIDS PARAMETERS

Parameter / Test grid	N	N_{Ramif}	$MinDepth_B$	$MaxDepth_B$
IEEE 33	33	3	3	12
IEEE 37	37	12	1	3

The comparison shows that the two test grids have a distinctive topology in terms of ramification and of depth of feeders, which brings variety in the study, particularly in the case of analyzing DSSE with GNN, which is also focused on the effect of branches between nodes.

V. RESULTS

While other studies are considering a small number of "Scenarios" regarding the renewables and storage resources placement correlated with a limited number of load value sets, in the tests presented in this paper it has been randomized all these parameters, such that most of the grid situations can be presented in the NN learning set. A number of 30'000 configurations of resources has been randomly generated for this purpose, in each of the two IEEE networks. These configurations have been split into training, validation and test sets with a rate of 70%, 10% and 20% respectively. All grid configurations have been calculated using OpenDSS as power flow tool [16,17].

The GNN implemented utilizes the attention mechanism for message passing between nodes, thus being a Graph Attention Network (GAT) [18]. A GNN message passing stream between operating instances was also implemented, with the model performing state estimation on 10 different operating instances in order to better ascertain the available resources in the network. The number of GAT layers used was set to 8, the hidden dimension to 512, and 16 independent heads were used. The loss function used for the masked modeling was chosen to be mean square error (MSE).

Moreover, both training and verification have been done with hiding factors between 50% and 95%. This factor represents the number of masked, thus unknown values resulting from the load-flow calculation and their inputs have been used for masking during training as well for the tests with scenarios used for validation. It is a big range of unknown data, chosen by purpose to test the state estimator degradation.

While 50% hidden values mean also 50% known values of the grids, 95% hidden values mean only 5% known values presented to the GNN based state estimator, so the range of degradation is from a basic situation of 50% known data to 10 times higher degradation, meaning only 5% known data.

In the Figure 3, showing the evolution of validation errors as applied to line currents, we observe that the NMAE rankings are increasing while the unknown values ratio is higher, thus being coherently with the masking rate.

We observe a graceful degradation in current error as masking error increases, which suggests that the model is able to utilize local context effectively for reconstruction.

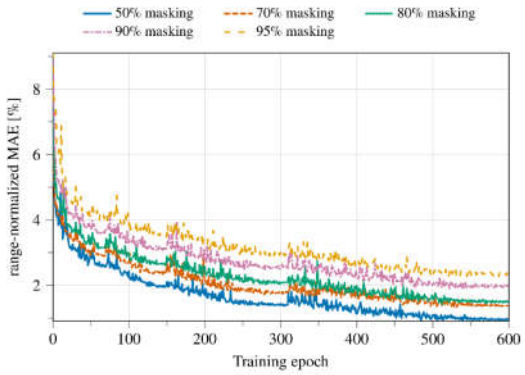


Fig. 3. Evolution of validation mean average error of currents for the IEEE 33 test grid.

In Figure 4 we observe that voltage estimation is more sensitive to masking rate variation than current is, with faster degradation once masking rates reach 95%. However, training does not appear to reach saturation for higher masking rates by epoch 600. This suggests that voltage magnitude estimation may more strongly depend on long-range context at high masking rates

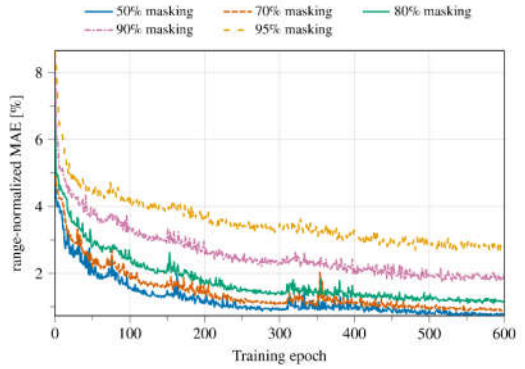


Fig. 4. Evolution of validation mean average error of voltages for the IEEE 33 test grid.

In Figure 5, we observe that the relative rankings of the models trained at different masking rates are mostly kept throughout training, with the notable exception of the 50% masking rate case learning slower, possibly due to reaching a local minimum, and of the 70% masking rate case surging over the 80% masking rate case for a while after epoch 300.

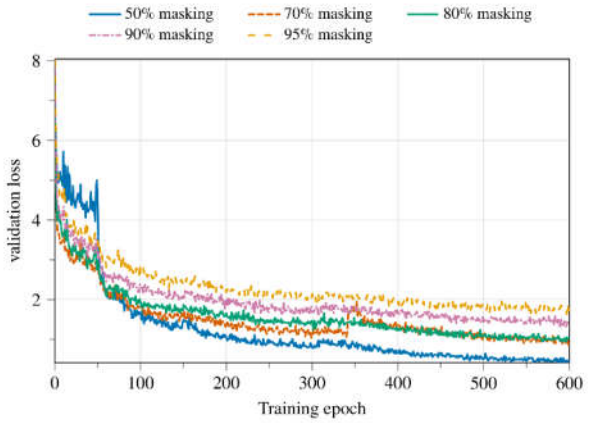


Fig. 5. Evolution of validation loss for the IEEE 33 test grid.

In Figures 6 and 7 we observe that the mean error constant across line active power magnitude (i.e. the estimator is homoscedastic), while it increases proportionally with line reactive power magnitude, revealing magnitude-dependent heteroscedasticity.

Meanwhile, Figure 8 reveals a broader range of errors for current estimation. There is a small degree of heteroscedasticity, mainly in smaller errors where currents are close to 0 and on the higher end of current values. For both active power and current estimation, we observe a small number of outliers with high errors from true current or power, but clustering around the same predicted current or power. This indicates a rare failure scenario where the model cannot gather enough information to estimate the real power or current, and instead collapses to a value close to the absolute mean of that estimated parameter.

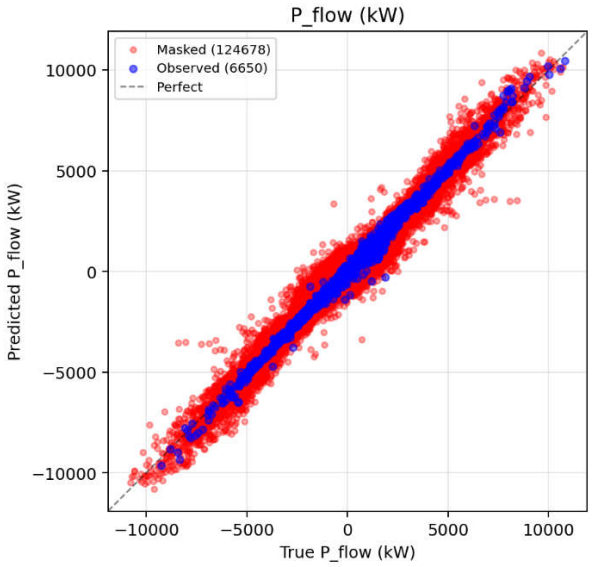


Fig. 6. Convergence chart of active power estimation for the IEEE 33 test grid.

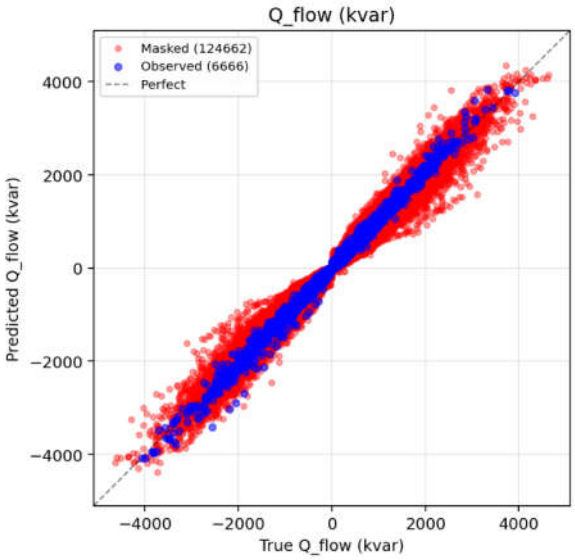


Fig. 7. Convergence chart of reactive power estimation for IEEE 33 test grid.

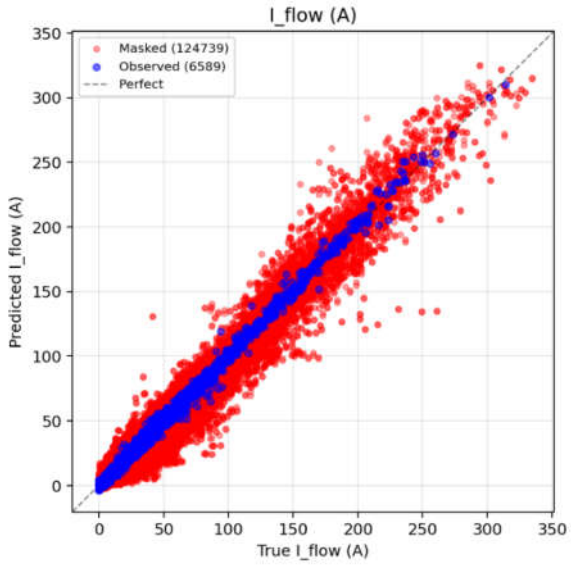


Fig. 8. Convergence chart of current estimation for IEEE 33 test grid.

Table II presents NMAE of different grid parameters of the IEEE test grid, for 50%, 70%, 80%, 90% and 95% hidden values.

TABLE II. TEST NMAE OF VARIOUS PARAMETERS AS MASKING RATE INCREASES FOR IEEE 33 TEST GRID

Parameter	***50%***	***70%***	***80%***	***90%***	***95%***
U	0.79%	0.95%	1.23%	1.90%	2.81%
U_{angle}	0.16%	0.16%	0.19%	0.27%	0.42%
P_{line}	0.95%	1.21%	1.26%	1.61%	1.91%
P_{bus}	1.29%	1.88%	1.82%	2.19%	2.55%
Q_{line}	0.43%	0.48%	0.51%	0.57%	0.67%
Q_{bus}	0.60%	0.70%	0.71%	0.79%	0.86%
I_{line}	0.98%	1.41%	1.51%	1.96%	2.35%
I_{bus}	1.52%	2.17%	2.56%	3.06%	3.47%

In the above table we observe the degradation of estimations as masking rate increases. We observe that voltage is the most affected, followed by active power. Reactive power is the least affected by increased masking rates, with mean average error increasing only by 50%. This may be due to stronger correlations between reactive powers and locally correlated powers, as well as more strongly constrained power factor priors in the generated dataset.

Figure 9 shows the graphical evolution of NMAE for the studied masking rates for the eight value types presented in the Table I. While at 50% masking rate all NMAE are practically up to 1.5%, for a 10 times degradation the highest NMAE was under 3.5%, which is a 2.33 times NMAE degradation for 95% comparing to 50% unknown values. This shows a high robustness of the GNN-based DSSE against a high number of missing input data.

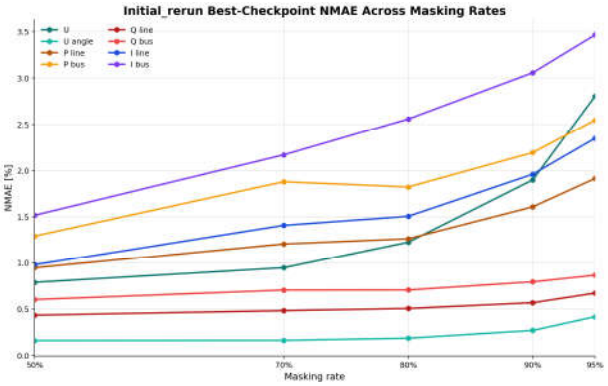


Fig. 9. Degradation of NMAE across masking rates for IEEE 33 test grid.

Same procedure has been applied to the IEEE 37 test grid. In the Figure 10 it is shown the evolution of validation errors for the IEEE 37 test grid.

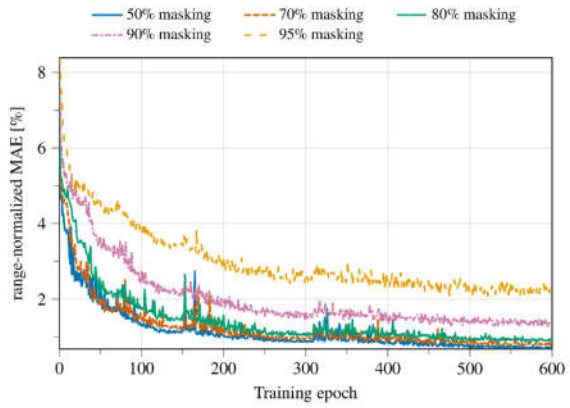


Fig. 10. Evolution of validation mean average error of voltages for the IEEE 37 test grid.

In Figure 11 we observe that validation loss when training on the IEEE 37 grid follows a more stable when compared to the IEEE 33 grid case. The model trained on the 50% masking rate case exhibits a similar overfitting behaviour in the first 50 epochs, after which the first warm restart forces it out of the local minimum. After this point, training follows a smooth improvement as epochs progress, indicating that the models do not begin overfitting by epoch 600.

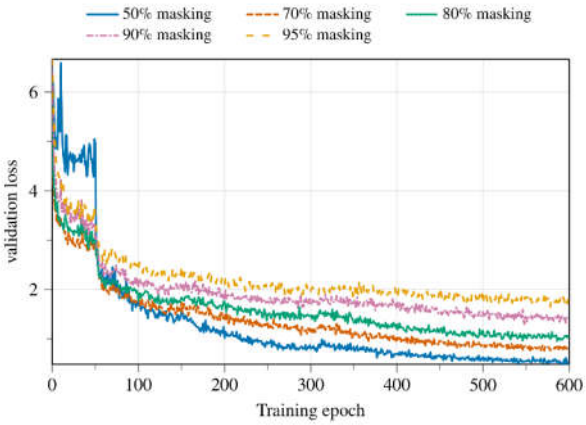


Fig. 11. Evolution of validation loss for the IEEE 37 test grid.

The test NMAE of different grid parameters of the IEEE test grid, for 50%, 70%, 80%, 90% and 95% hidden values, using the final checkpoint of each model, is presented in Table III.

TABLE III. TEST NMAE OF VARIOUS PARAMETERS AS MASKING RATE INCREASES FOR IEEE 37 TEST GRID

Parameter	***50%***	***70%***	***80%***	***90%***	***95%***
U	0.76%	0.87%	0.98%	1.49%	2.33%
U_{angle}	0.19%	0.18%	0.20%	0.26%	0.37%
P_{line}	0.63%	0.76%	0.85%	1.08%	1.25%
P_{bus}	1.13%	1.43%	1.61%	1.96%	2.23%
Q_{line}	0.29%	0.31%	0.32%	0.39%	0.45%
Q_{bus}	0.50%	0.57%	0.62%	0.69%	0.75%
I_{line}	0.73%	0.96%	1.11%	1.38%	1.58%
I_{bus}	1.32%	1.91%	2.36%	2.88%	3.29%

In Table III we can observe the evolution of error variance with regards to masking rate. We observe that the variance of estimation of bus currents has a high variance to mean error ratio, suggesting that a smaller set of outliers is contributing heavily to variance. We observe that for most variables, the ratio of variance to mean average error remains relatively stable, in some cases mildly decreasing as in the case of bus currents. This suggests that while outliers, such as in particularly isolated buses, contribute heavily to mean error, their effect is not sensitive to high masking rates.

For IEEE 37 test grid, at 50% masking rate all NMAE values are under 1.4%, and even with 10 times fewer observed values the highest NMAE was under 3.3%, which is a 2.35 times degradation in NMAE for 95% unknown values compared to the 50% case.

Additionally, we observe in both tables that degradation is graceful up to 80%, but sharply increases with the higher masking rate regimes.

VI. CONCLUSIONS AND NEXT WORK

This work has analyzed the degradation of a GNN-based DSSE outputs for 50% up to 95% hidden values, in order to assess error rates achievable for a wide range of unknown inputs from the grid. In order to better understand the degree

to which state estimation degrades, two different medium voltage test grids have been used, namely IEEE 33 and IEEE 37. While having a similar number of nodes, the topology of these test grids has different levels of ramifications and number of feeders between them, highlighted by the three topology parameters introduced in section IV. The case of 50% hidden vales has been used as baseline for both test grids. A GNN has been trained and tested on 30000 different examples for each IEEE test grid, assessing NMAE for several grid parameter.

For both the IEEE 33 and IEEE 37 test grids, at 50% masking rate the NMAE for all parameters is under 1.5%, while for a 10 times degradation in observed values the highest NMAE was under 3.5% - with a slightly better performance for IEEE 37 grid. The NMAE degradation for 95% compared to 50% unknown values has a maximum ratio of 2.35, while the unknown values have been 10 times more. This highlights the robustness of the GNN-based DSSE against a high number of missing input data.

Future work will analyze the degradation of NMAE for increased rates of unknown input data in the case of additional NN training methods, such as Physics Informed Neural Network (PINN) reinforced training. Moreover, additional test grids will study the influence of measurement distortion on the results and will investigate the placement of measurements in the extreme cases.

ACKNOWLEDGMENT

This work has received funding from the Digital Europe Programme (DIGITAL) EU funding programme, under grant agreement No 101158494 - ECLIPSE DIGITAL project (Energy Consumption Reduction based on Open-source Reference framework) and also was supported by a grant of the Ministry of Research, Innovation and Digitalization - UEFISCDI, project number PN-IV-P8-8.2-EUD-2025-0062, within PNCDI IV. This text reflects only the author's views, and the Commission is not liable for any use that may be made of the information contained therein.

REFERENCES

[1] F. Ahmad, A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş, "Distribution system state estimation-A step towards smart grid," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 2659–2671, Jan. 2018, doi: https://doi.org/10.1016/j.rser.2017.06.071

[2] Sruthi Ankam; Vedik Basetti; Sachidananda Sen, Advanced Distribution System State Estimation Techniques: A Review, 2025 IEEE 1st International Conference on Smart and Sustainable Developments in Electrical Engineering (SSDEE), 28 February 2025 - 02 March 2025, Dhanbad, India

[3] Jinfeng Hu; Shang Ma; Xundong Gong; Jiakai Ling; Jun Qin, Distribution System State Estimation Based on Multi-section Smart Meter Data, 2021 IEEE Sustainable Power and Energy Conference (iSPEC), 23-25 December 2021, Nanjing, China

[4] Zui Tian; Haoxiang Fang, Three-Phase State Estimation of Multi-Source Data Based on Real-Time Monitoring of Active Distribution Networks, 2026 International Conference on Digital Analysis and Processing, Intelligent Computation (DAPIC), 26-28 February 2026, Incheon City, Korea

[5] Shaofang Wang; GuangYi Liu, State estimation for active distribution networks basedon multisource hybrid measurements, 2016 China International Conference on Electricity Distribution (CICED), 10-13 August 2016, Xi'an, China

[6] Sreedevi Valsan Kandenkavil; Nitin Padmanbahan, Performance Analysis of Optimization Based Static Distribution State Estimation Techniques, 2020 International Conference on Power, Instrumentation, Control and Computing (PICC), 17-19 December 2020, Thrissur, India

[7] Qian Qiang; Sun Guoqiang; Xu Wei; Yan Minghui; Wei Zhinong; Zang Haixiang, Distribution System State Estimation Based on Pseudo Measurement Modeling Using Convolutional Neural Network, 2018 China International Conference on Electricity Distribution (CICED), 17-19 September 2018, Tianjin, China

[8] Huiheng Yi; Zhenyuan Zhang; Peng Wang, State Estimation of Distribution Network with the Improved Deep Residual Neural Network, 2022 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), 08-11 July 2022, Shanghai, China

[9] Huayi Wu; Zhao Xu; Minghao Wang, Unrolled Spatiotemporal Graph Convolutional Network for Distribution System State Estimation and Forecasting, IEEE Transactions on Sustainable Energy (Volume: 14, Issue: 1, January 2023)

[10] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention Is All You Need." In Advances in Neural Information Processing Systems 30. 2017.

[11] IEEE 33 test network for medium voltage stuies, https://www.torontomu.ca/content/dam/cue/research/reports/33bus%20test%20system.pdf, accessed on 10.06.2026

[12] Pegado, R.; Ñaupari, Z.; Molina, Y.; Castillo, C. Radial distribution network reconfiguration for power losses reduction based on improved selective BPSO. Electr. Power Syst. Res. 2019, 169, 206–213

[13] IEEE PES Test Feeder, https://cmte.ieee.org/pes-testfeeders/resources, accessed on 20.06.2026

[14] K. P. Schneider, B. A. Mather, B. C. Pal, C. W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. de Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott, and W Kersting, "Analytic Considerations and Design Basis for the IEEE Distribution Test Feeders," IEEE Transactions on Power Systems, Volume: 33, Issue: 3, May 2018.

[15] W. H. Kersting, "Radial distribution test feeders," in Proc. IEEE Power Eng. Soc. Winter Meeting, pp. 908–912, Jan. 2001.

[16] OpenDSS Code, EPRI Distribution System Simulator, https://sourceforge.net/p/electricdss/code/HEAD/tree/trunk/Distrib/IEEETestCases/37Bus/, accessed 10.06.2026

[17] OpenDSS - Electric power Distribution System Simulator (DSS), https://www.epri.com/pages/sa/opendss, accessed on 15.03.2026

[18] Z. Ren, X. Chu, and H. Ye, "State Estimation for Reconfigurable Distribution Systems Using Graph Attention Network," *IEEE Access*, vol. 11, pp. 107237–107250, Jan. 2023, doi: https://doi.org/10.1109/access.2023.3316017.