Machine Learning Based State Estimation for Transmission and Distribution Grids

Anamitra Pal (Email: Anamitra.Pal@asu.edu)
Assistant Professor
Arizona State University (ASU), USA

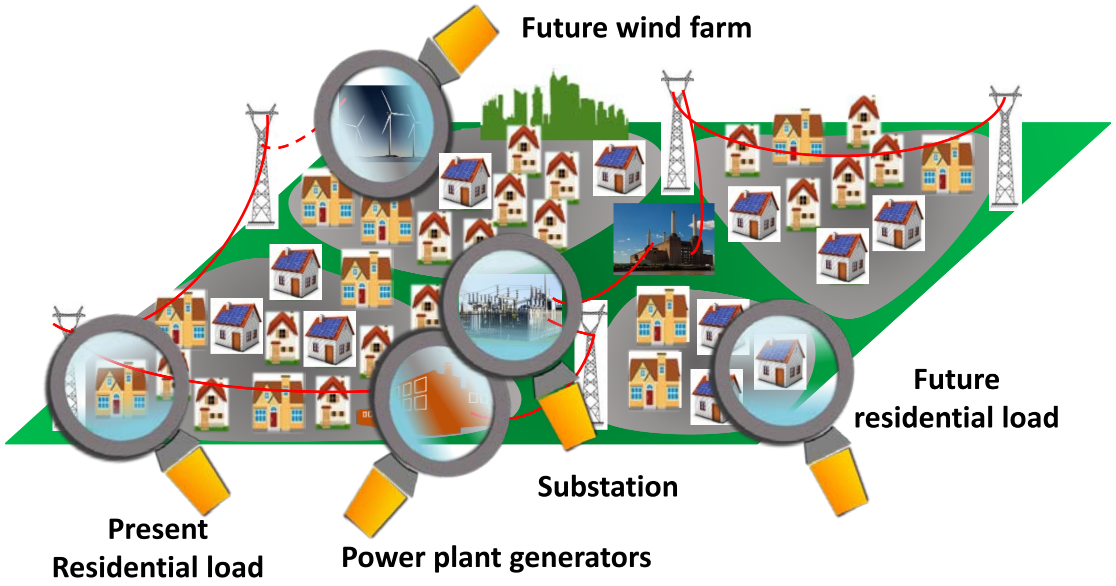
Evangelos Farantatos (Email: efarantatos@epri.com)
Principal Project Manager
Grid Operations & Planning
EPRI, USA

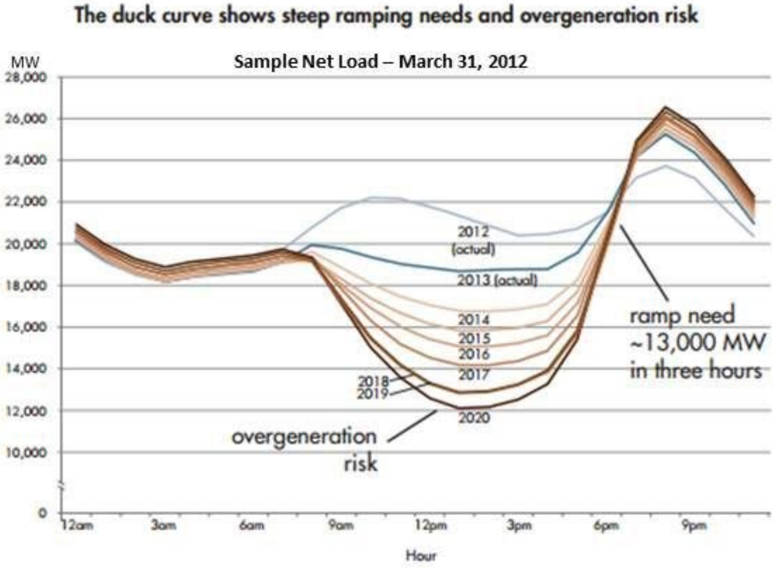
Acknowledgements: This work was supported in part by the Department of Energy under grants DE-AR00001858-1631 and DE-EE0009355, the National Science Foundation (NSF) under grant ECCS-2145063, and the Electric Power Research Institute (EPRI) under grant 10013085.

Outline

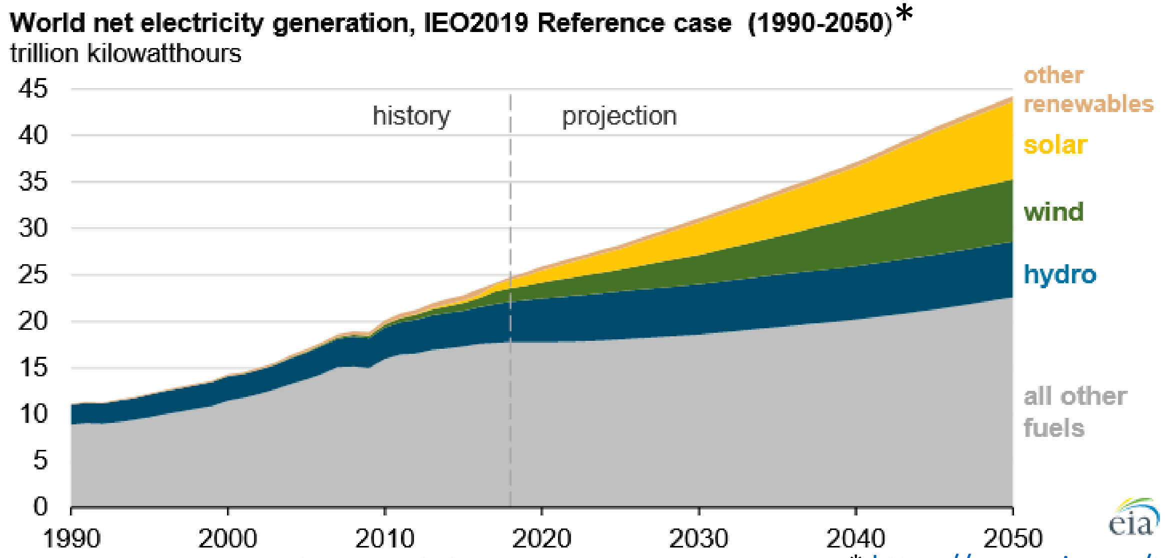
- Introduction
- State Estimation Background
- Machine Learning Based State Estimation
- Application in Distribution Systems
- Application in Transmission Systems
- Handling Topology Changes and Bad Data
- Summary & Future Work

Power Systems … Now and in the Future





(from the California Independent System Operator)



Source: U.S. Energy Information Administration, *International Energy Outlook 2019** https://www.eia.gov/todayinenergy/detail.php?id=41533

Characteristics of Sensors Continuously Monitoring Electric Power Grid

	Transmission		Distribution		
	SCADA	PMU	Smart Meter	SCADA	µPMU/ D-PMU
Spatial Resolution	Very dense	Becoming dense	Dense	Sparse	Extremely sparse
Temporal Resolution	1 – 5 seconds	**< 33 milliseconds**	1 min – 1 hour	1 – 5 seconds	**< 16 milliseconds**
Latency	2 – 4 seconds	**< 1 millisecond**	Few hours to days	2 – 4 seconds	**< 50 milliseconds**
Time-synchronized?	No	**Yes**	No	No	**Yes**

- Time-synchronized measurements will continue to play an important role in the high-speed, high-precision monitoring, protection, and control of modern power systems!!!

SCADA: Supervisory control and data acquisition PMU: Phasor measurement unit µPMU: Micro-PMU D-PMU: Distribution-PMU

Conventional State Estimation

Backbone EMS Function for Situational Awareness

- **State Definition [x]**: Positive sequence voltage phasors (bus voltage magnitudes and angles) of system’s buses
- **Measurement Set [z]**: From SCADA system
  - Voltage magnitude, current magnitude, real & reactive power flows and injections
  - Measurement model: Nonlinear $[z]=[h([x])]+[e]$
  - Gaussian distribution of measurement error
- **Solution Algorithm**: Weighted Least Squares
  - Iterative Solution

Biased State Estimation

- **Unbalanced Grid Operation/System Asymmetries**
- **Measurement Time Skewness**
- **Large Scale Problem/Long Execution Time (30 secs-3 mins)**

State Estimation & Synchrophasor Technology

Unbalanced Operation & System Asymmetries

Three-phase measurements
Three-phase formulated State Estimation

PMU

Measurement Time Skewness

GPS Synchronized and Time Tagged Phasor Measurements

Large Scale Problem – Computationally Demanding

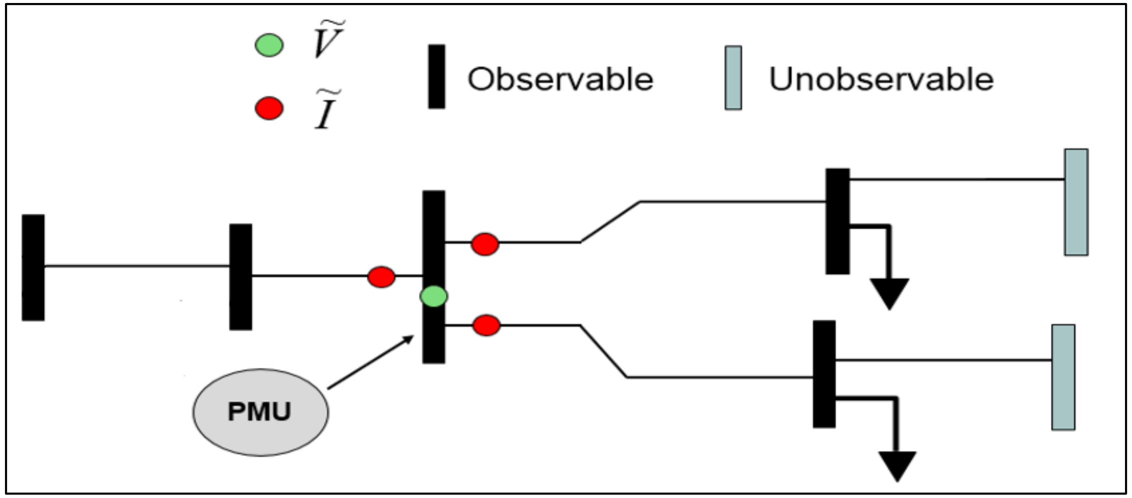
Linear State Estimation

Linear State Estimation

PMU Measurement Based State Estimation

- **State Definition [x]**: The same
- **Measurement Set [z]**: From PMUs
  - Voltage and current phasors
  - Measurement model: Linear $Z = H(X) + e$
  - Gaussian distribution of measurement error
- **Solution Algorithm**: Weighted Least Squares
  - Direct Solution

PMU Observability



Drawback: Requirement for full system observability from PMUs

Can machine learning (ML) help overcome the disadvantages of a purely synchrophasor-based state estimation approach???

Least Squares Approach vs. Bayesian Approach

Least squares estimation:

$$\hat{x}_{\text{WLS}}(z) = \arg\min_{x} \|z - h(x)\|^2$$

Minimizes the modeling error

Minimum mean squared error (MMSE) for Bayesian estimation:

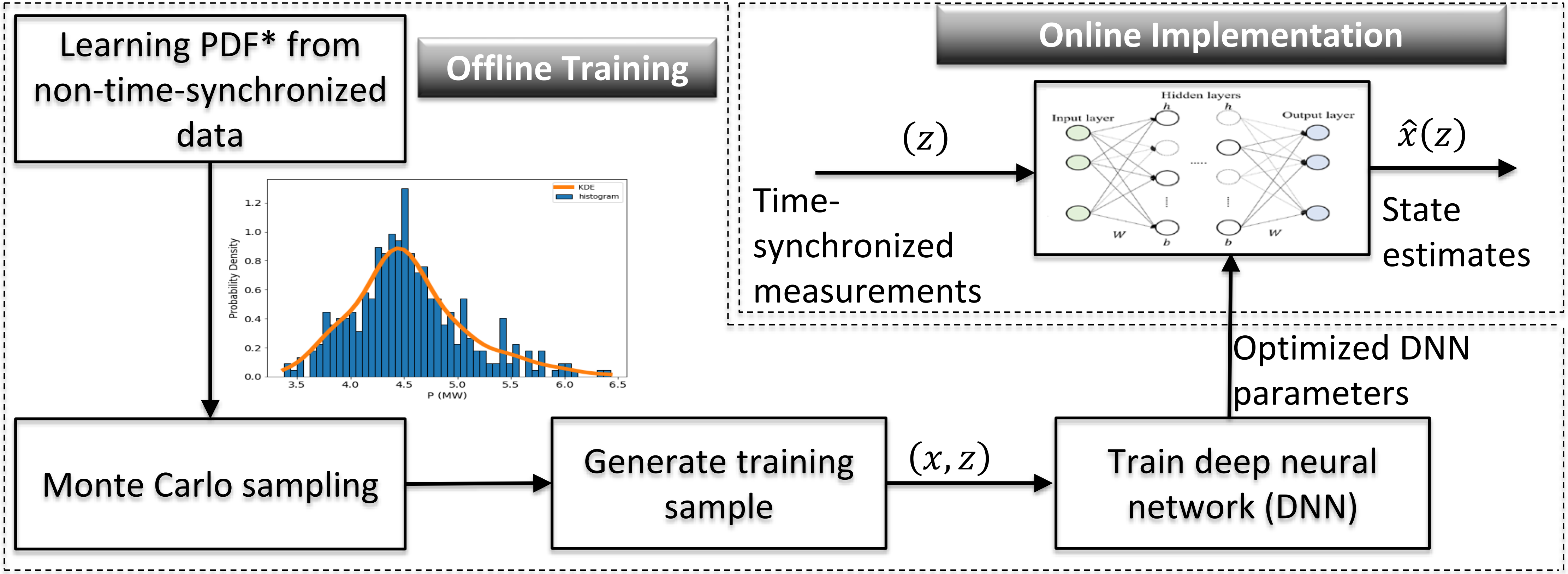
$$\min_{\hat{x}(.)} \mathbb{E}(\|x - \hat{x}(z)\|^2) \Rightarrow \hat{x}^*(z) = \mathbb{E}(x|z)$$

Directly minimizes the estimation error

- Computation of the conditional mean can be exceedingly complex
- The underlying joint distribution of x and z is unknown or impossible to specify, making the direct computation of $\hat{x}^*$ intractable
- Deep neural network (DNN) can be used to approximate the MMSE state estimator

x: state variable (voltage phasors)
z: measurement

Schematic of the Proposed Methodology



*PDF: probability density function

- Non-time-synchronized data is only used to generate sample data to train the DNN (Offline operation)
- Time-synchronized measurements are used in the testing stage (Online operation)
- Does not require complete system observability by PMUs/μPMUs/D-PMUs

Least Squares-based State Estimation vs. Proposed DNN-based State Estimation[1]

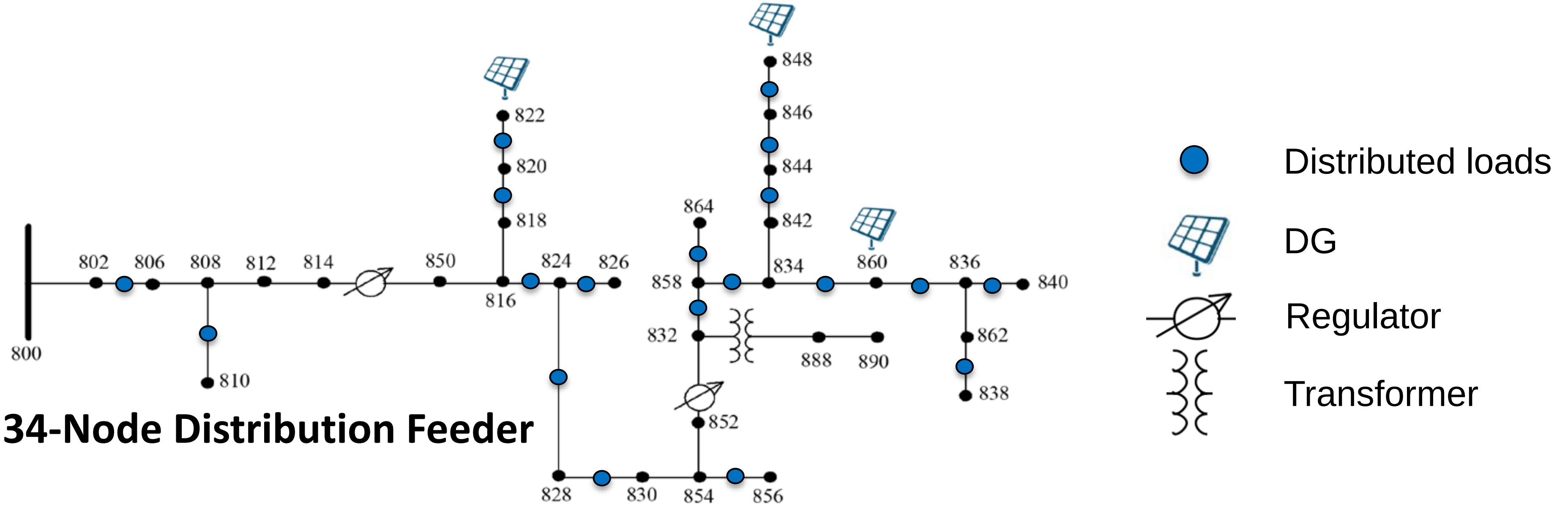
Classical Least Squares	Deep Neural Network
Increased observability is required for better online operation	Observability is not needed for online operation
Slow online computation time	Fast online computation time
Susceptible to non-Gaussian noise in the measurements	Relatively immune to measurement noise characteristics
Synchronizing inputs from different sensors is a challenge	Issue of incompatible timescales is mitigated (non-time-synchronized data is not directly used in state estimation)
Accuracy is a function of observability	Accuracy is a function of training
Simple to solve	Parameter tuning is difficult

[1] B. Azimian, R. S. Biswas, A. Pal, and L. Tong, "Time synchronized distribution system state estimation for incompletely observed systems using deep learning and realistic measurement noise," in *Proc. IEEE Power Eng. Soc. General Meeting*, Washington DC, pp. 1-5, 26-29 Jul. 2021.

Characteristics of Distribution System and DNN

- **Distribution System Characteristics:**
  - Single, double and three phase unbalanced Wye-Delta loads
  - ZIP load models
  - Voltage regulators and transformers
  - Single, double, and three phase laterals
  - Distributed generation (DG)

- **Deep Neural Network (DNN) Characteristics[2]:**
  - 5 Hidden layers – 200 neurons/layer
  - ReLU activation function for hidden layers
  - Linear activation function for output layer
  - Optimizer: ADAM
  - Empirical mean squared error loss function
  - Separate neurons and layers for each phase



IEEE 34-Node Distribution Feeder

[2] B. Azimian, R. S. Biswas, S. Moshtagh, A. Pal, L. Tong, and G. Dasarathy, "State and topology estimation for unobservable distribution systems using deep neural networks," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1-14, Apr. 2022.

µPMU/D-PMU Placement to facilitate DNN-based Distribution System State Estimation (DSSE)

- Conventional optimal µPMU/D-PMU placement strategies aim for complete system observability
- New algorithm is proposed that enables DNN-based DSSE by exploiting the correlations that exist between the input features

Algorithm: µPMU/D-PMU Placement for DNN-based DSSE

Inputs: Budget, $\text{DSSE}_{\text{accuracy}}$, $\text{Corr}_{\text{Threshold}}$, Number of nodes $= M$

Output: Location of the µPMUs/D-PMUs

A.i. $N_{cluster} = 1$

A.ii. Calculate correlation coefficient between each voltage phasor V_{ij}^{kl} $\forall i \in \{A, B, C\}, \forall j \in \{\text{mag}, \text{ang}\}, \& \ \forall k, l \in \{1, \dots, M\}$

A.iii. If correlation coefficient $\forall k, l \in \{1, \dots, M\}$ is greater than $\text{Corr}_{\text{Threshold}}$ for $\forall i \in \{A, B, C\}$ & $\forall j \in \{\text{mag}, \text{ang}\}$ then go to (A.vii.)

A.iv. $N_{cluster} = N_{cluster} + 1$

A.v. Cluster each correlation coefficient matrix for $\forall i \in \{A, B, C\}$ & $\forall j \in \{\text{mag}, \text{ang}\}$

A.vi. Find common node in each cluster for each correlation coefficient and place µPMU/D-PMU on this node.

A.vii. If $\text{DSSE}_{\text{accuracy}}$ is satisfied or µPMU/D-PMU installation cost $\geq$ Budget, then End, else go to (A.iv.)

Generating Training Database for IEEE-34 Node Distribution System

Mapping Data
- To generate realistic loading scenarios using Pecan Street data

Perform KDE
- To fit distribution to loads

Do Power Flows
- To find voltages and currents

Add Noise
- To create realistic measurements

Type of Data	#Samples
Training	10,000 (80%)
Testing	2,500 (20%)

KDE: Kernel density estimation

DNN-based State Estimation for IEEE-34 Node Distribution System

Results for Linear State Estimation (LSE) and Proposed DNN-based State Estimation*

Method	Phase MAE (degrees)	Magnitude MAPE (%)	#Nodes
Linear State Estimation (LSE)	0.0194	0.0352	22[3]
DNN-based State Estimation	0.0241	0.0386	3

* When measurement noise is Gaussian

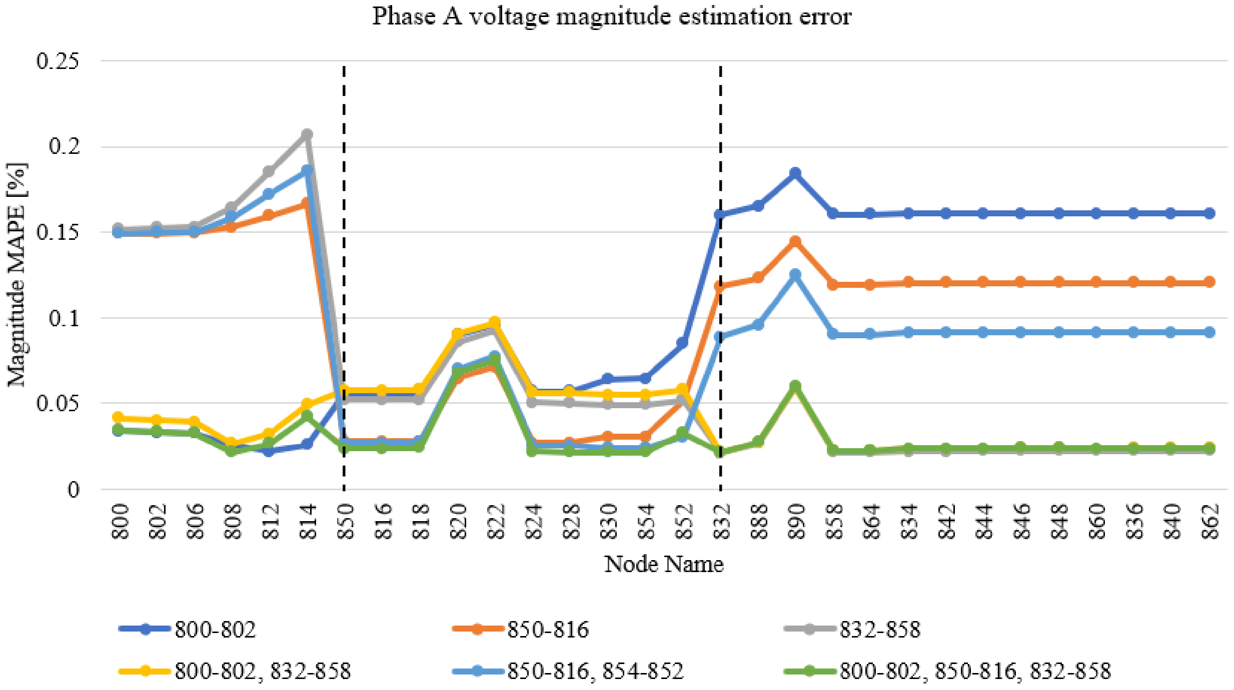
DNN-based State Estimation Results Under Gaussian and Non-Gaussian Measurement Noise

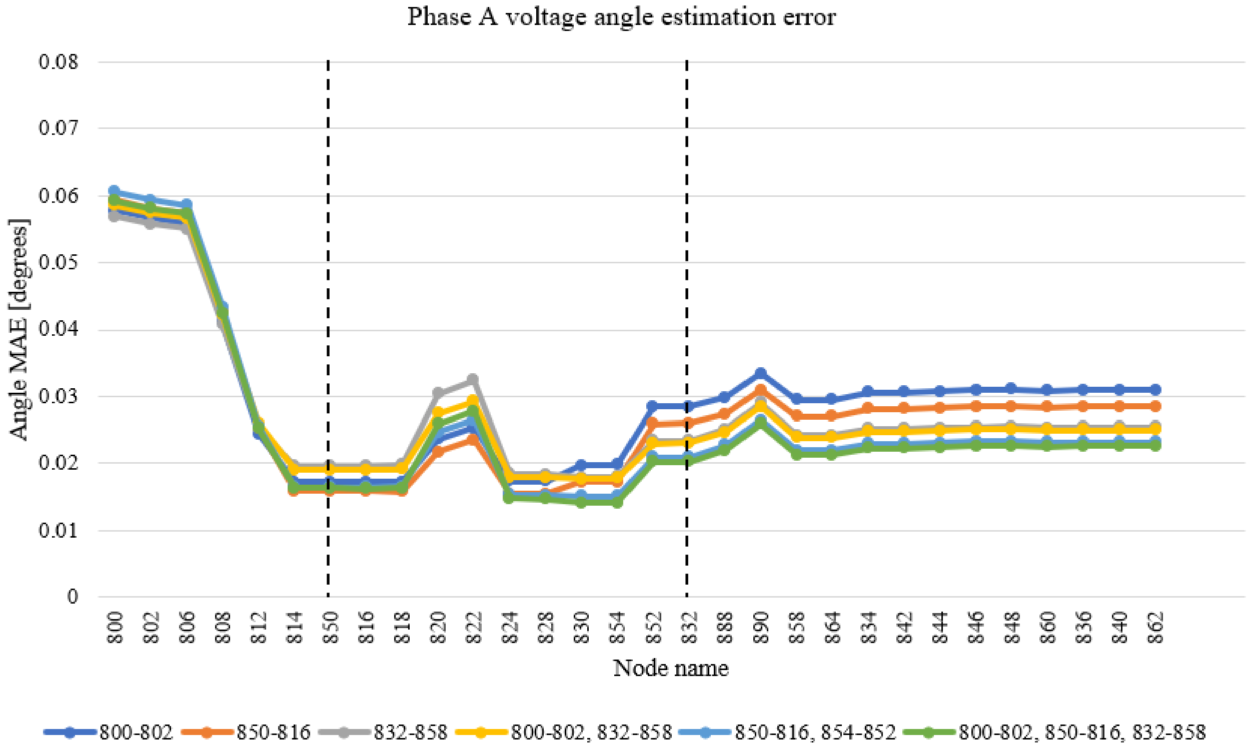
Method	Phase MAE (degrees)	Magnitude MAPE (%)	#Nodes
DNN with Gaussian Noise	0.0241	0.0386	3
DNN with Non-Gaussian Noise	0.0242	0.0393	3

$$MAPE = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{y_j - \hat{y}_j}{y_j}\right| \qquad MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right|$$

[3] R. S. Biswas, B. Azimian, and A. Pal, "A micro-PMU placement scheme for distribution systems considering practical constraints," in *Proc. IEEE Power Eng. Soc. General Meeting*, Montreal, Canada, pp. 1-5, 2-6 Aug. 2020.

Demonstrating Impact of µPMU/D-PMU Placement on DNN-based DSSE Performance



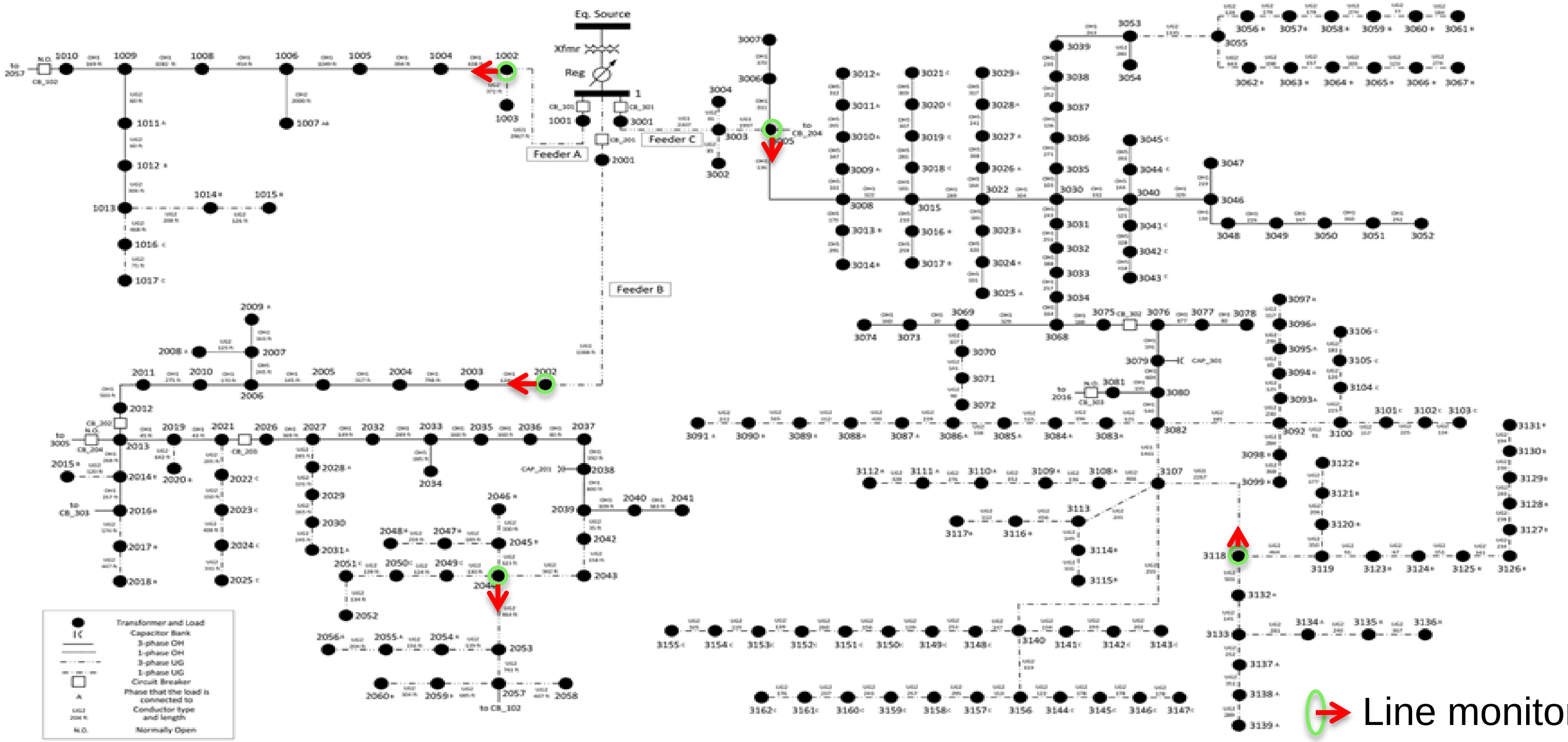


- Minimum magnitude estimation error was obtained when the three µPMUs/D-PMUs were placed in three different clusters
- µPMU/D-PMU placement does not significantly influence angle estimation error as the intercorrelations are very high

Demonstrating High-Speed Tracking Ability of the DNN-based State Estimator

- A trained DNN performs a matrix multiplication of the input values with the weights and biases of its neurons – a process that can be executed very fast
- The DNN took only 0.01 seconds to produce the estimates

DNN-based State Estimation for 240-Node U.S. Midwest Distribution System



Line monitored by µPMU/D-PMU

- System Characteristics[4]:
  - 3 feeders, 240 nodes
  - One-year worth of hourly smart meter data available
  - Overhead lines, under-ground cables, capacitor banks, LTC transformers, line switches, and secondary distribution transformers
  - OpenDSS model available

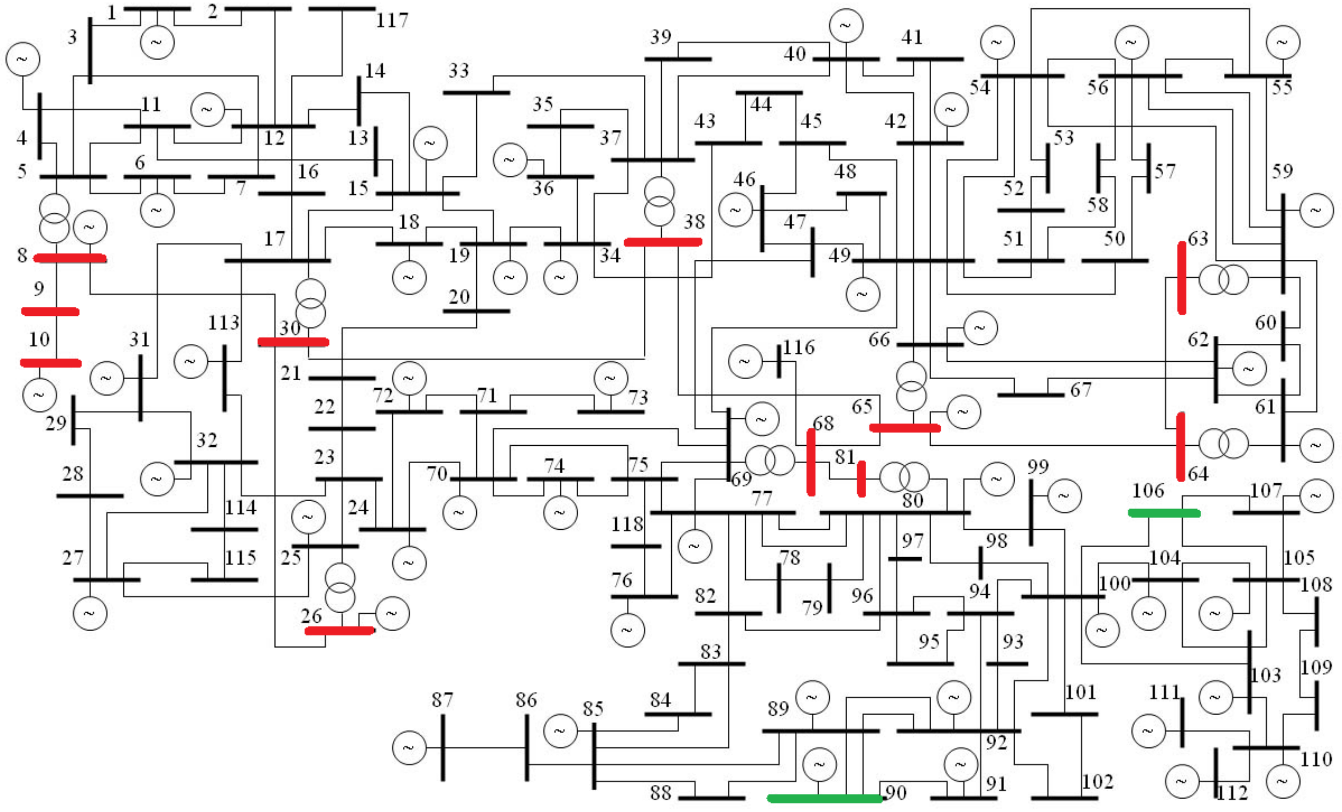
Method	Phase MAE (degrees)	Magnitude MAPE (%)	#Nodes
DNN-based DSSE	0.0081	0.0144	5

[4] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, "A Time-Series Distribution Test System Based on Real Utility Data," *2019 North American Power Symposium (NAPS)*, Wichita, KS, USA, 2019, pp. 1-6

DNN-based State Estimation for Transmission Systems – Incremental PMU Placement

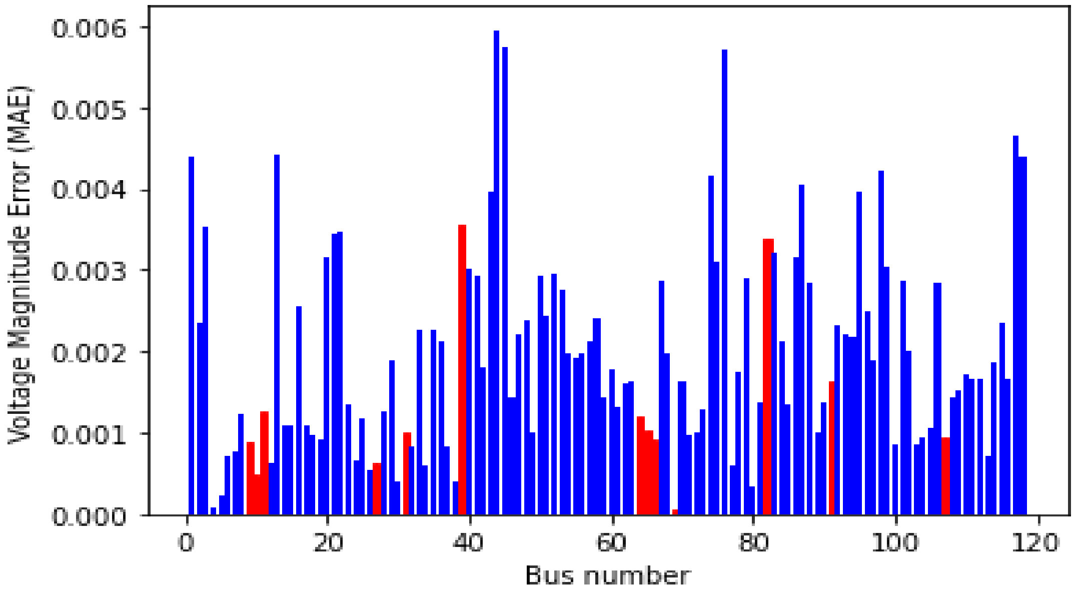
- New PMUs added based on two criteria:
  - High variability in training dataset
  - Distance from existing PMUs

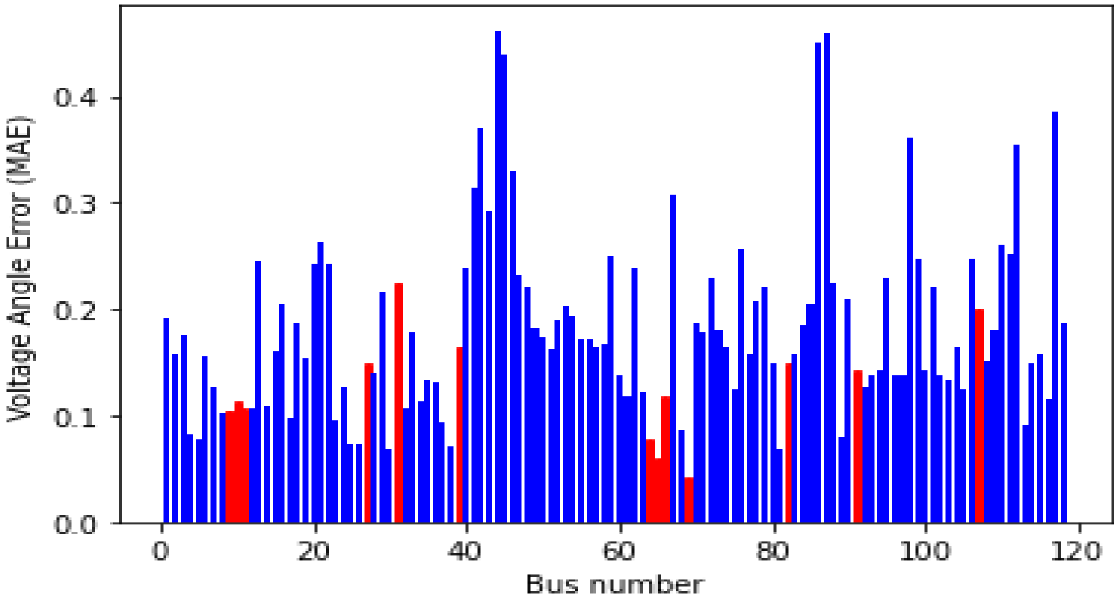
Method	Phase MAE (°)	Magnitude MAPE (%)	#Buses
LSE with Gaussian Noise	0.1693	0.9051	32[5]
DNN-SE with Gaussian Noise	0.1453	0.1209	13
DNN-SE with Laplacian Noise	0.1528	0.1579	13



[5] A. Pal, G. A. Sanchez, V. A. Centeno, and J. S. Thorp, "A PMU placement scheme ensuring real-time monitoring of critical buses of the network," *IEEE Trans. Power Del.*, vol. 29, no. 2, pp. 510-517, Apr. 2014.

DNN-based State Estimation for IEEE 118-Bus Transmission System (per bus view)





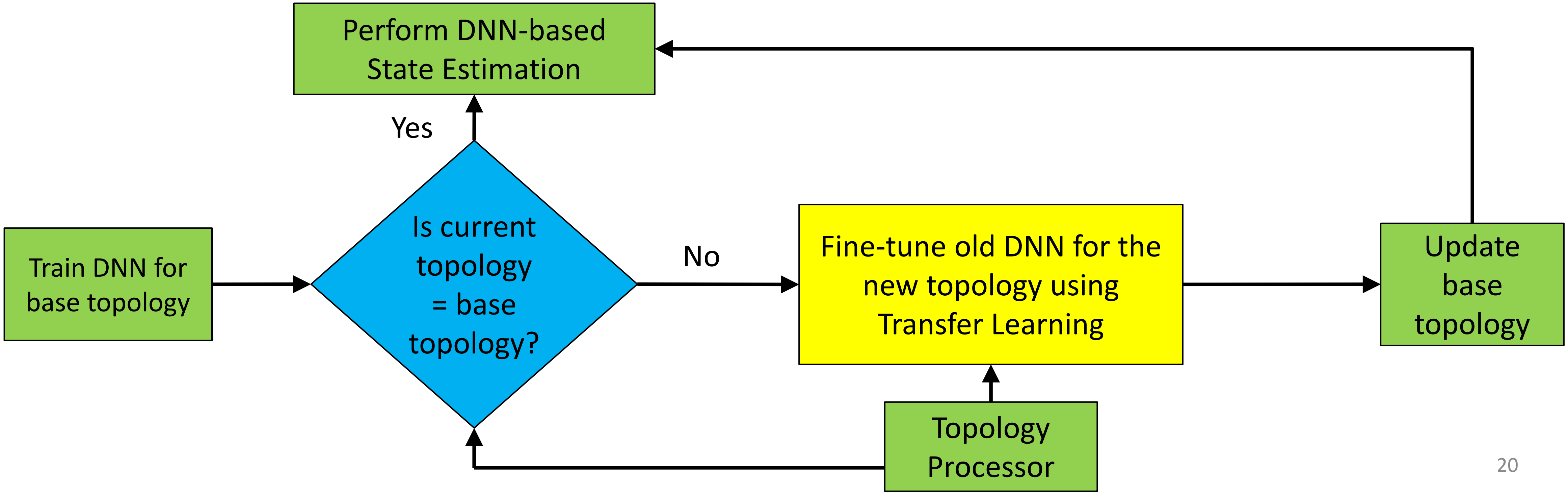
Red color: Bus locations where PMUs are placed

Blue color: Bus locations where PMUs are not placed

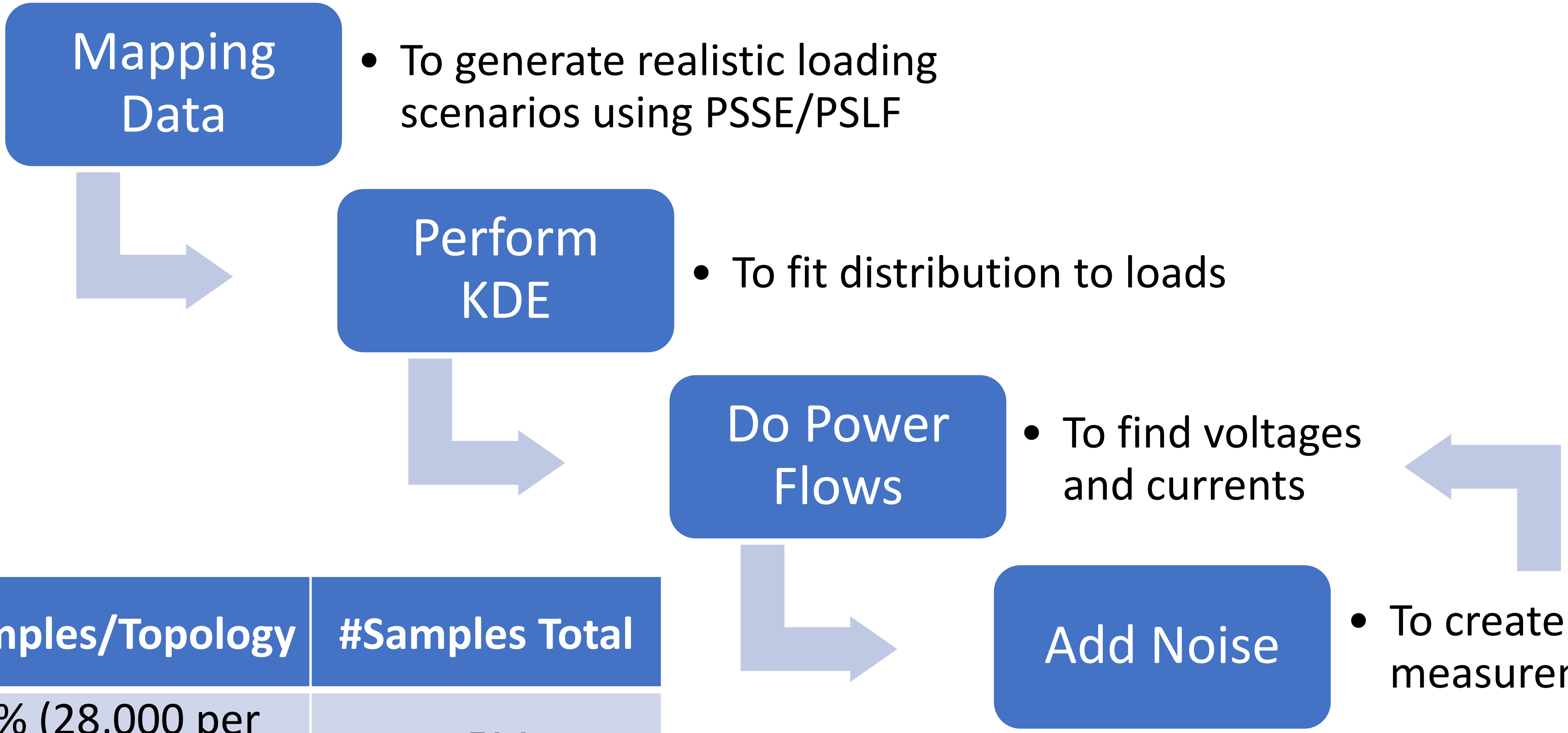
Evaluation Metric	Value
Phase MAE (°)	0.1453
Magnitude MAPE (%)	0.1209

DNN-based State Estimation Under Varying Topologies

- So far, the DNN-based state estimation (DNN-SE) was trained for a given (fixed) topology
- However, if the DNN is tested with different topologies, its performance can deteriorate
- Transfer Learning is the ability to fine-tune a DNN's parameters for a given change in training and testing environment



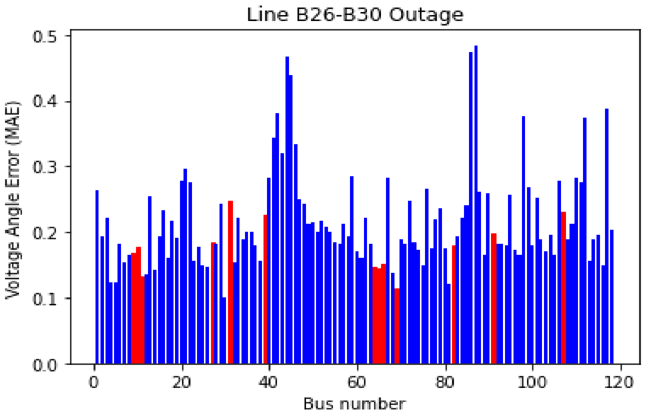
Generating Training Database for IEEE 118-Bus Transmission System for Handling Topology Changes

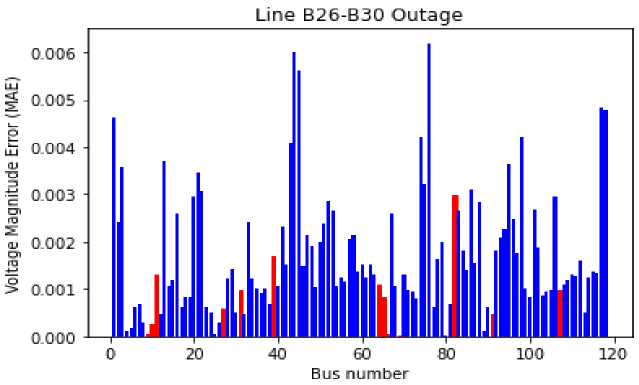


- To generate realistic loading scenarios using PSSE/PSLF
- To fit distribution to loads
- To find voltages and currents
- Repeat for different topologies
- To create realistic measurements

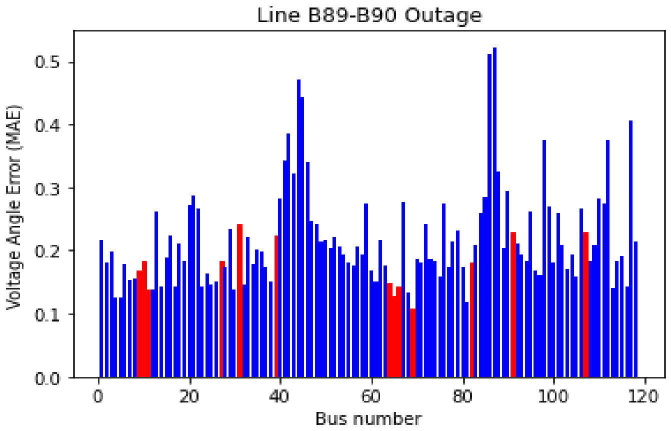
Type of Data	#Samples/Topology	#Samples Total
Training	80% (28,000 per topology)	≈5M
Testing	20% (4,000 per topology)	≈700K

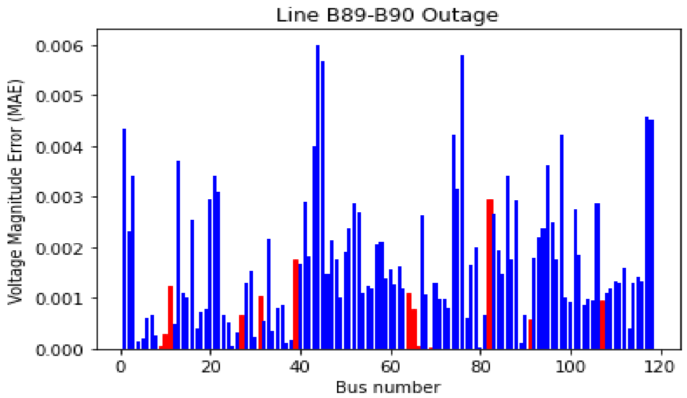
Transfer Learning Results for Different Topologies of the IEEE 118-Bus System



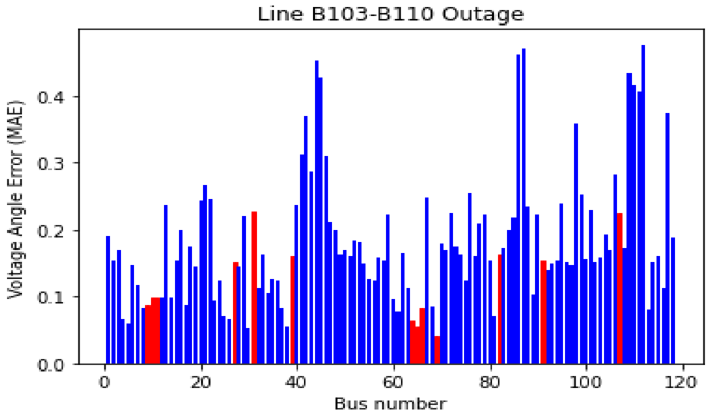


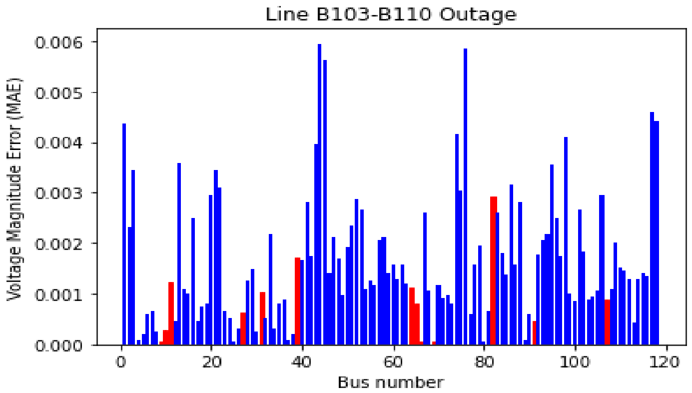
Evaluation Metric	Value
Phase MAE (°)	0.1544
Magnitude MAPE (%)	0.1279





Evaluation Metric	Value
Phase MAE (°)	0.1502
Magnitude MAPE (%)	0.1268





Evaluation Metric	Value
Phase MAE (°)	0.1490
Magnitude MAPE (%)	0.1258

Fine-tuning took ≈ 30 seconds

Bad Data Detection for the DNN-based State Estimator

- Due to the unobservability problem, conventional bad data detection and correction approaches are not suitable for this DNN-based state estimator
- The use of Wald Test to detect bad data has been proposed previously[6]

$\mathcal{H}_0$: without bad data
$\mathcal{H}_1$: with bad data
z: μ_0, σ_0^2

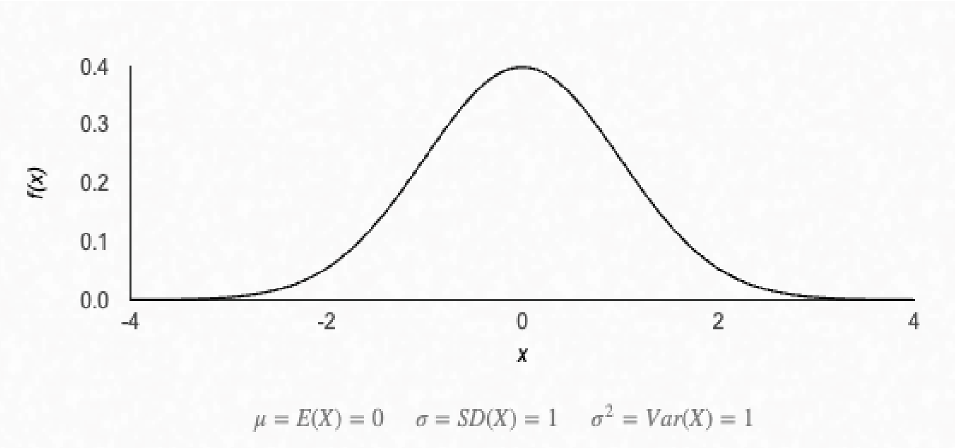
α Wald Test

$$\left|\frac{z-\mu_0}{\sigma_0}\right| \gtrless Q^{-1}(\alpha/2)$$

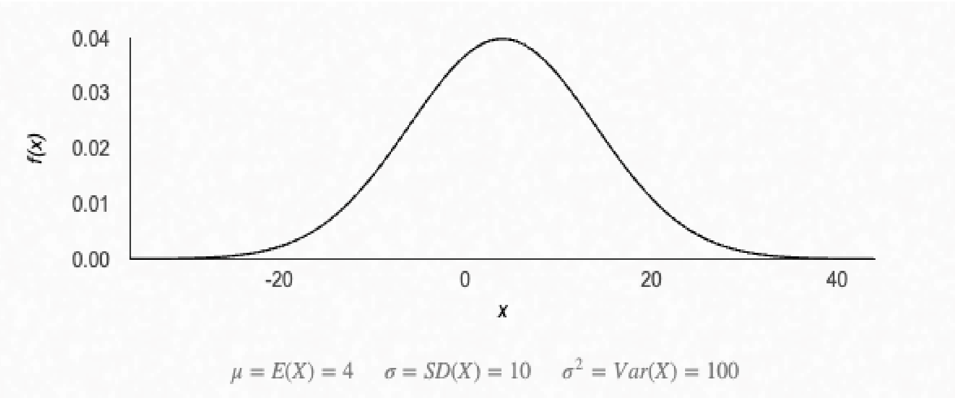
where, $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{u^2}{2}\right) du$

- Bad data is detected when the deviation exceeds the threshold set by α

Normalized Input Data



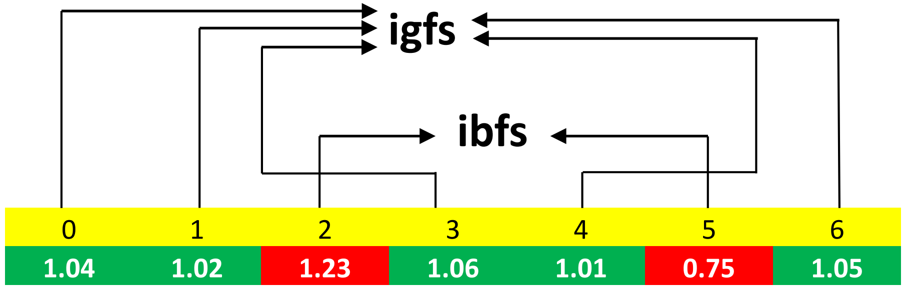
Bad Data



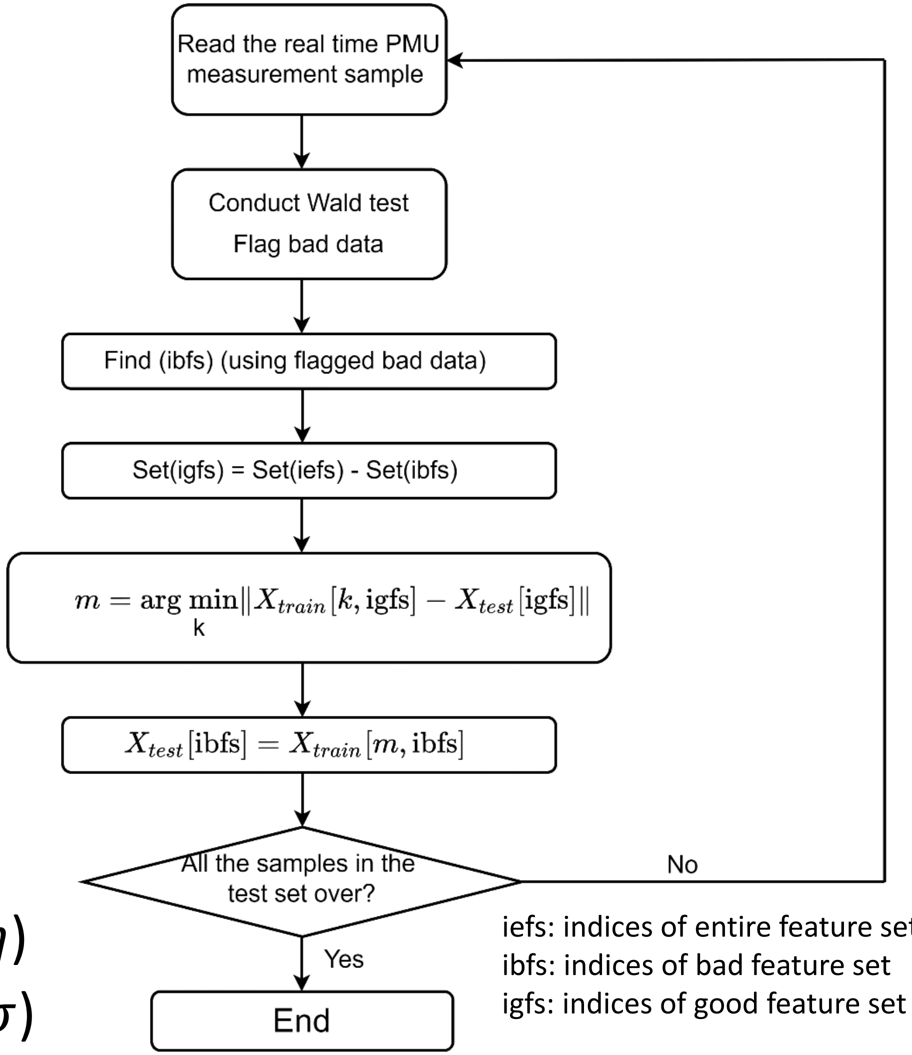
[6] K. R. Mestav, J. Luengo-Rozas and L. Tong, "Bayesian State Estimation for Unobservable Distribution Systems via Deep Learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910-4920, Nov. 2019.

Bad Data Correction for the DNN-based State Estimator

- α is typically set at 0.05, implying that the false alarm (false positive) probability is no greater than 5%
- Contrary to replacing the bad data by its corresponding training data mean[6], it is replaced by the operating condition (OC) that is closest in the training database



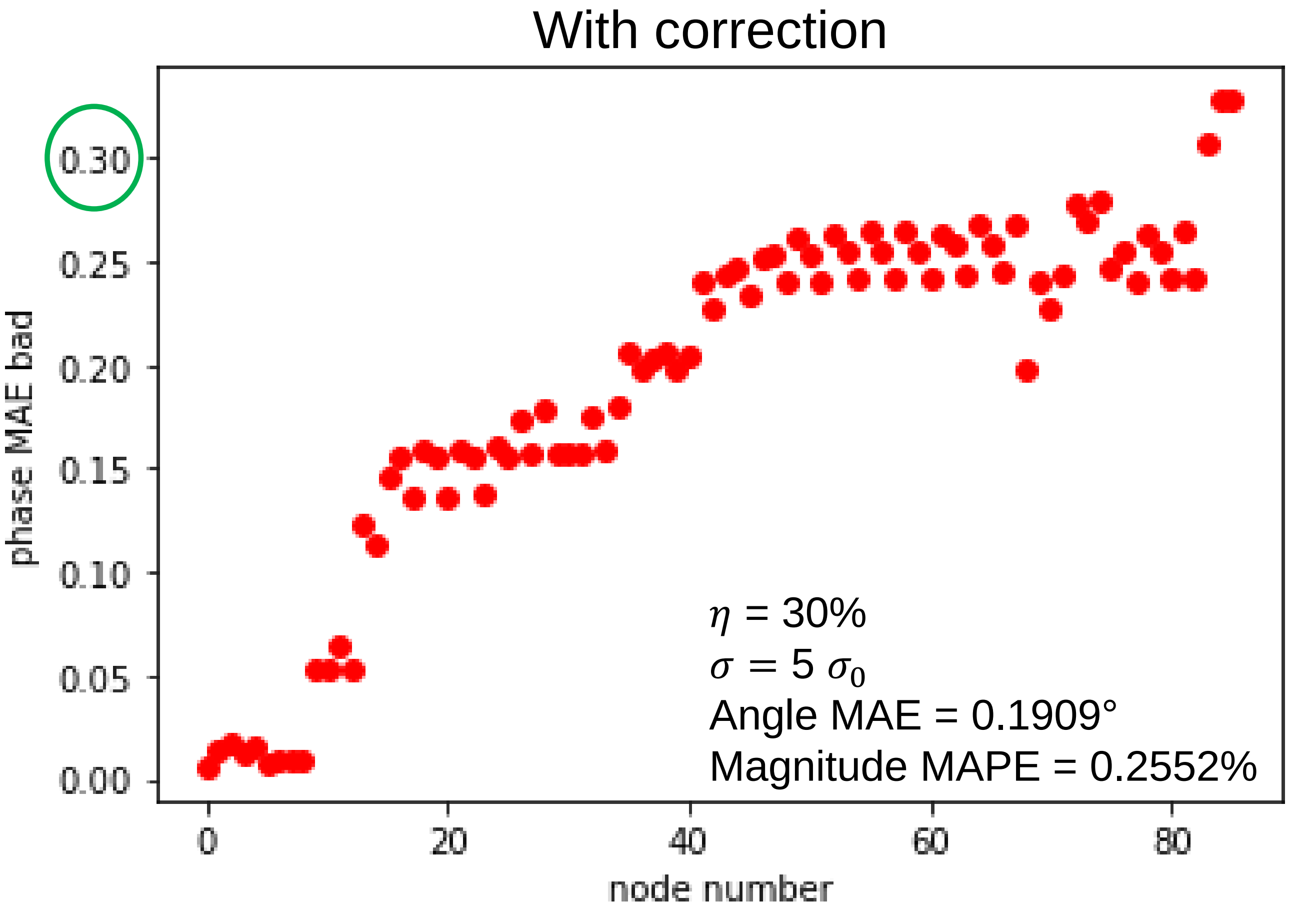
- Furthermore, two types of bad data are investigated:
  - Amount of bad data (expressed in terms of variations in η)
  - Badness of bad data (expressed in terms of variations in σ)

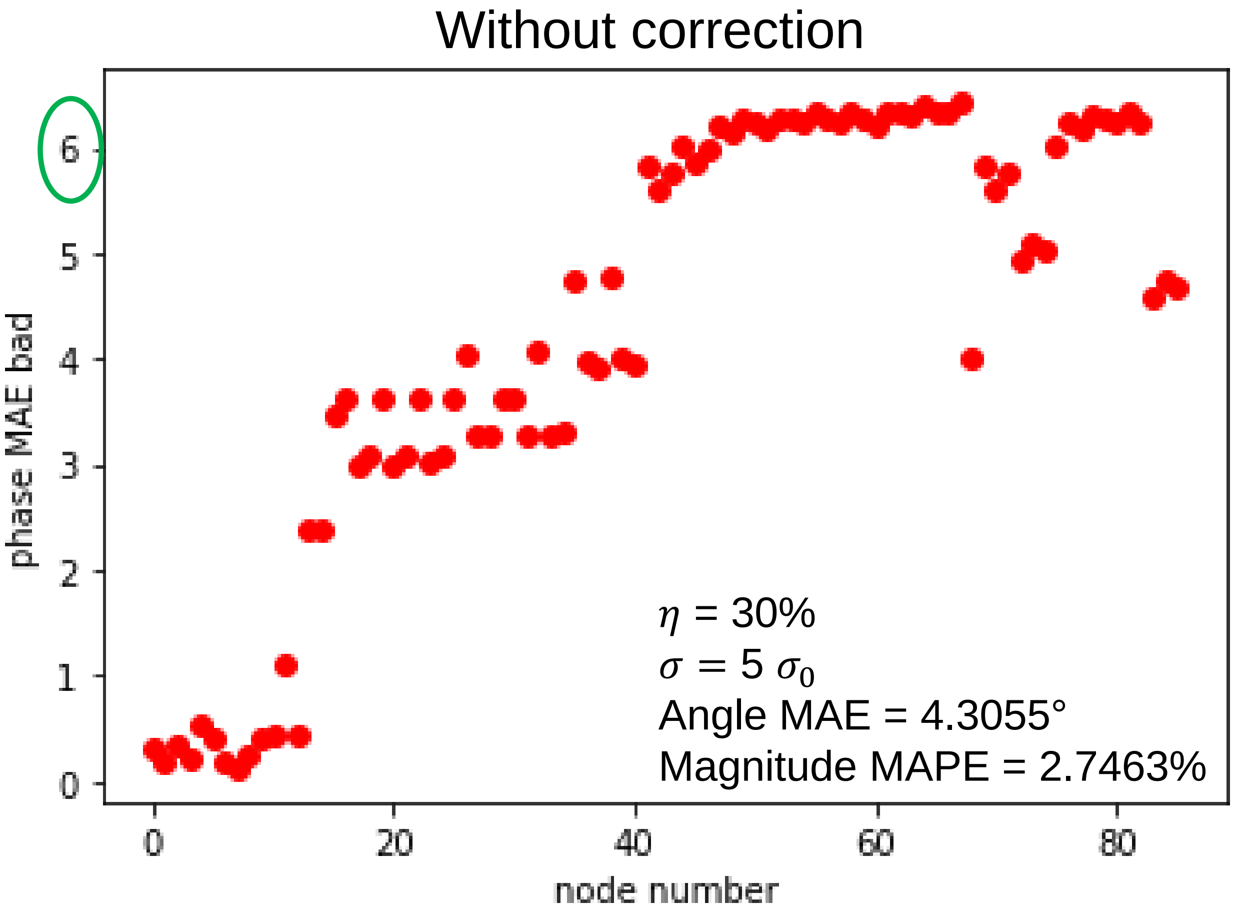


iefs: indices of entire feature set
ibfs: indices of bad feature set
igfs: indices of good feature set

[6] K. R. Mestav, J. Luengo-Rozas and L. Tong, "Bayesian State Estimation for Unobservable Distribution Systems via Deep Learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910-4920, Nov. 2019.

Impact of Bad Data Correction on DNN-based State Estimation Accuracy (Distribution)

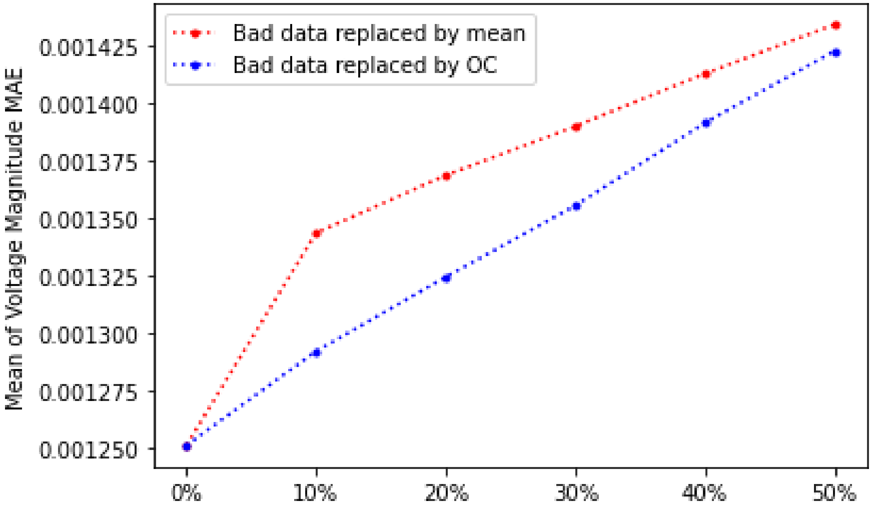


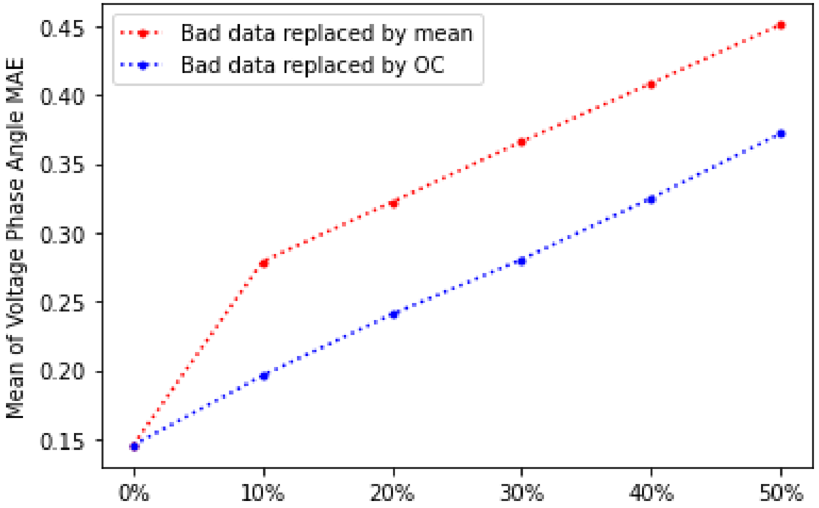


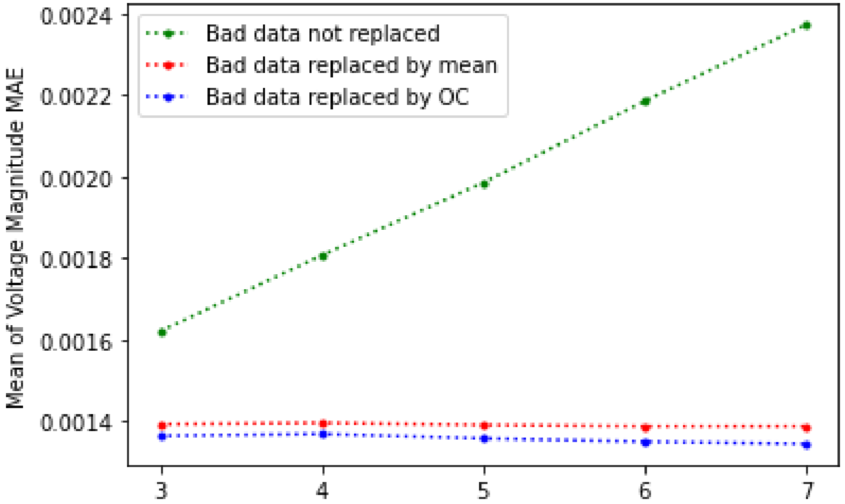
IEEE-34 Node Distribution Feeder

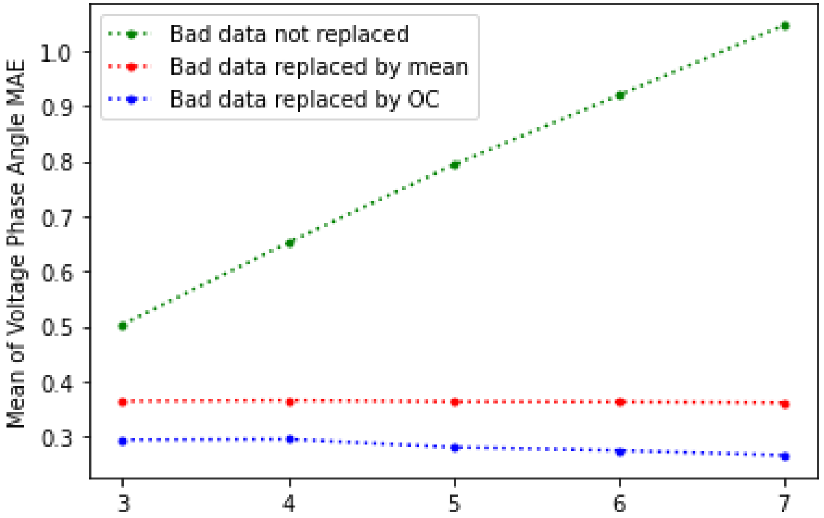
Impact of Bad Data Correction on DNN-based State Estimation Accuracy (Transmission)

Variation in amount of bad data (η)









Variation in badness of bad data (σ)

- Replacing bad data using the nearest OC results in higher accuracy than replacing using the mean value

Summary and Future Scope of Work

- A methodology to perform time-synchronized state estimation using deep learning was formulated for systems that are incompletely observed by PMUs/μPMUs/D-PMUs
- Different strategies to place PMUs/μPMUs/D-PMUs for improving the state estimator's performance were explored
- Ability to handle topology changes and bad data were demonstrated
- Ongoing work:
  - Providing robustness guarantees to DNN performance
  - Incorporating physics of the system during training data generation
  - Developing advanced monitoring, protection, and control capabilities using the obtained insights

Antos Varghese

Behrouz Azimian

Satya Sahoo

Shiva Moshtagh

Hritik Shah