8-31-2021

Application of Artificial Neural Networks to Power System State Estimation

James P. Carmichael
University of Kentucky, james.carmichael@uky.edu

Yuan Liao
University of Kentucky, yuan.liao@uky.edu

Application of Artificial Neural Networks to Power System State Estimation

James P. Carmichael, University of Kentucky, Yuan Liao, University of Kentucky, USA

Contact author: Yuan Liao, yuan.liao@uky.edu, USA

Abstract

State estimation function is essential for effective and timely execution of power system automation and control systems, especially in modern active distribution systems where more intermittent renewable energy systems are integrated into the grid. Distribution system state estimation faces a lot of challenges including lack of monitoring devices and possible incorrect topology information. Developing efficient state estimation for distribution systems is thus of great interest. This paper presents results on utilizing artificial neural networks for this purpose.

Artificial neural networks have been used in power distribution system state estimation. However, there is a lack of systematic analysis and study of which types of ANNs and what structures including parameters are most suitable for state estimation applications. When designing an ANN for a state estimator, *trial and error approach* has been common and there is no systematic method available to guide the process. The ultimate goal of the research is to examine the performance of various types of ANNs (e.g., Multi-Layer Perceptron (MLPs), Convolutional Neural Networks (CNNs) and Long-Short-Term-Memory Networks (LSTMs)) with different structures and also provide possible guidance on how to choose the different parameters, including model parameters such as number of hidden layers and number neurons in a layer, and algorithm parameters such as adjustable learning rate, for desired performance metrics. The paper presents preliminary results based on MLPs. IEEE standard 34-bus test system is used to illustrate the proposed methods and their effectiveness.

The paper seeks to contribute to a more systematic approach to neural network and deep learning applied to power system state estimation, thus enhancing situational awareness, system resiliency and real-time monitoring and control of power distribution systems. Successful state estimation function will increase the ability of distribution systems to integrate more renewable energy based generations.

Keywords—artificial neural networks (ANNs), multilayer perceptron networks (MLPs), convolutional neural networks (CNNS), long short-term memory networks (LSTMs), distribution system state estimation (DSSE)

I. INTRODUCTION

In power systems an essential requirement is that of *resiliency*. In general, resiliency includes the ability of a power system to withstand and recover quickly from events that may be considered low-frequency, yet high-impact events or adverse conditions.

Examples of such events or adverse conditions relate to but are not limited to the following: Extreme weather, Natural disasters, Man-made outages (physical, cyber, coordinated), Lack of Observability,Topology Errors, and False Data Injection Attacks (FDIA).

State estimation process provides optimal estimate of the true values of bus voltages and angles and power flows across the power system [1][2]. The results provide the basis or enhancement for other power system applications such as system planning, optimization, fault analysis, protection, and fault location [3][4][5][6].

This paper focuses on application of artificial neural networks to *distribution system state estimation (DSSE)* and will investigate the ability of such networks to ensure resiliency to any events that may compromise data integrity. There are virous types of networks such as Conventional Feed-Forward Multi-Layer Perceptron Networks (MLPs) / Deep Neural Networks, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) / Long Short-Term Memory Networks (LSTMs), and Hybrid-Neural Networks Utilizing a Combination of Network Types. Preliminary results based on MLPs are presented in this paper.

II. BACKGROUND

A. *Review of Conventional State Estimation*

State estimation research and application has historically been largely focused on *transmission systems* as opposed to *distribution systems*. With increasing developments of the “smart grid”, increased utilization of phasor measurement units (PMUs) and improvements in monitoring and communications, *Distribution System State Estimation (DSSE)* interest and research has greatly increased in recent years.

The inherent challenges of application of “conventional” state estimation techniques to power distribution systems based upon *weighted least squares* is well established in the literature.

In recent years, “deep learning neural networks” have gained increasing interest in not only being able to improve the weighted least squares method, but also in the possibility of being able to address what may be considered as “extreme” or “adverse” conditions such as, but not limited to *lack of observability, topology errors, false data injection attacks*, network outages due to weather or malicious attack, and variances in weather that may affect distributed power generation from solar and wind sources.

Conventional state estimation was introduced in 1970 via a series of papers authored by Fred C. Shweppe and J. Wildes. The overall problem, mathematical modeling and general algorithm for state estimation, error detection and identification are presented in [1].

The key assumption of the *classical approach* presented is that the state estimation vector consisting of the voltage magnitude and phase angles at all generation and load buses is *static* or *quasi- static*. Further assumptions are that the system is *balanced*, linear and can be accurately approximated via an iterative algorithm utilizing *weighted least squares* as the estimator. While these assumptions are reasonable when applied to transmission systems, they may not hold for distribution systems.

An approximate model and the resulting simplifications in state estimation, bad data detection and identification are presented in [2]. This model is based on a DC load flow yielding linear equations with the following four basic assumptions:

- Reactance over resistance of all lines are significantly larger than one
- Magnitude of voltage ≈ 1 for all buses
- Angle differences between voltages at two ends of all lines are close to zero
- Existence of errors in real power measurements

The resulting approximate model, while enabling potential application to distribution systems is not readily applicable to state estimation in general for practical transmission or distribution networks. Thus, [7] addresses implementation problems associated with dimensionality, computational efficiency, data storage and the time-varying nature of actual power systems.

The time-variation inherent in power systems is addressed in [8]. This paper is a review of *dynamic state estimation (DSE)* methods as opposed to *static state estimation (SSE)*. These methods are based primarily on Kalman Filtering (KF) techniques, M-estimation, and the Square Root Filter (SRF) technique which is an alternative implementation of KF that is numerically more stable.

Paper [9] discusses the essential role of power system *observability* to the state estimation problem and presents a theoretical basis for an algorithm to determine observability. The authors emphasize the requirement that conventional or classical state estimation methods be applied only to systems that are observable and thus establish that an observability test be conducted prior to performing state estimation. The algorithm presented is based upon a *graph theoretical* or *topological approach*. Specifically, the algorithm seeks to determine if the Jacobian of the system parameter network h(x) is full rank. If so, the power system network is considered observable.

The challenges to state estimation due to lack of observability are further discussed in [10]. The authors reiterate the essential observability criteria needed in order to perform classic state estimation and further surmise that the first step to *controllability* is observability.

Again, the challenges imposed by the dynamic nature of power systems and especially that of distribution systems with high penetration of distributed energy resources (DERs) is noted as a significant barrier to the application of classical state estimation techniques.

While the authors do recognize the improvements that the placement of smart meters and PMUs have made in enhanced situational awareness and greater observability, they also point out that smart meters do not offer real-time updates and that the practical implementation of PMUs is and will continue to be limited due to their cost.

In the paper being referenced, *robustness* refers to the insensitivity of the state estimation algorithm to major deviations in a limited number of redundant measurements. Thus, it is clear that the challenges of applying classical state estimation methods based upon weighted least squares and similar estimators to distribution systems also extend to determination of system controllability, observability and robustness.

The authors in [11] provide an in depth discussion of the growing threats to modern power system resiliency that applies to all aspects of the grid (i.e. generation, transmission, distribution, distributed generation, micro-grids, etc.). Investment in the modernization of the power grid must be done so with a "No Regrets Strategy". This strategy is based upon the *cornerstones* of *resiliency*, *flexibility* and *connectivity*.

- Resiliency – Resistance to High-Impact, Low Frequency Events
  - Extreme Weather
  - Earthquakes, Tsunamis
  - Man-made Outages (Physical, Cyber, Coordinated)
- Flexibility – Adaptability to Uncertainties
  - Fuel Prices
  - Power Market Prices/Incentives
  - Variable Generation
  - Consumer Behavior
  - Regulation and Policy
- Connectivity – Enhanced Interoperability Across Electricity Enterprise
  - Advanced Sensors
  - Mobile Devices
  - Grid Modernization
  - Two-Way Flow

B. Distribution System State Estimation (DSSE)

State estimation was first introduced by Fred C. Schweppe and J. Wildes in 1970 for power systems. States are defined as the vector of the voltage magnitudes and angles at all network buses [1]. Novel approaches on system resource scheduling considering reserve were presented in [3] [4], and advanced methods for protection and fault locations for distribution systems were described in [5][6]. These techniques can all benefit from improved measurements and topology.

Essentially, state estimation algorithms provide for a means of eliminating or minimizing measurement noises and errors and possible topology errors that would otherwise prevent accurate determination of the system state values at all buses. Power system state estimation was initially introduced and applied to transmission systems only and then extended to distribution systems, considering substantial differences between distribution and transmission systems.

Among these differences are the radial topology, low X/R ratios, phase imbalances and relative lack of measurement devices inherent in distribution systems [2]. With the emergence of the smart grid and distributed generation (DG), such as photovoltaic systems, wind turbines, electric vehicle to grid (V2G) technology and other forms of power penetration, power flow is now bi-directional as opposed to previously being unidirectional.

Additionally, given the unpredictable nature of renewable energy sources such as solar and wind energy as well as the varying real-time utilization of power inherent in distributed networks, updated state estimation algorithms is now necessary.

As mentioned previously, challenges to the application of "conventional" state estimation as applied to distribution systems relate directly to the fundamental differences of the two power system types.

Figure 1 presents an example of both types of power networks and some of the differences that pose a challenge to the direct application of conventional state estimation to distribution systems.

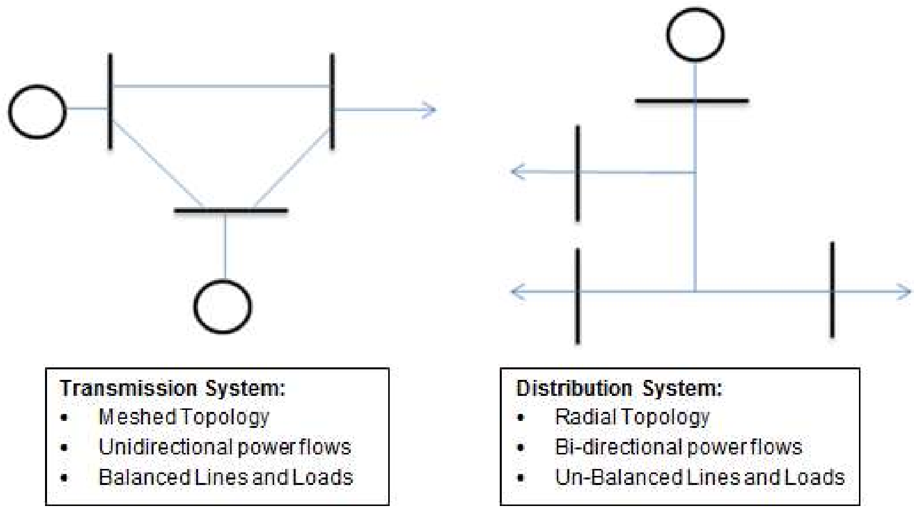

Figure 1 - Transmission and Distribution System Key Characteristics

To appreciate the challenges that the emerging smart distribution grid pose to the direct application of conventional state estimation, it is essential to first understand the inputs and functional blocks that enable state estimation. Figure 2 provides an overview of the inputs and main functional blocks.

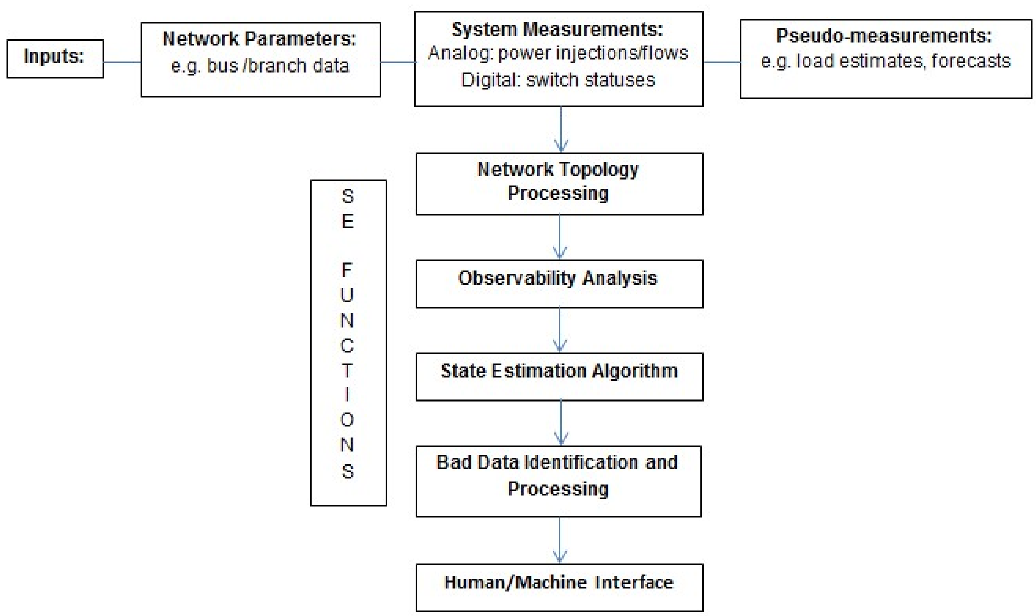

Figure 2 - Functional Block Diagram of State Estimation

Note that the *Network Topology Processing* functional block verifies the accuracy of the network parameters included as *Inputs*. The *Observability Analysis* functional block establishes that there is sufficient data available for the *State Estimation Algorithm* functional block, and these two blocks may be integrated together in some methods. As discussed earlier, the relative lack of metering in distribution networks reduces the "observability" of the system.

The ability to meet this challenge, while being improved through the implementation of "smart meters" such as PMUs (phasor measurement units), will continue to be an inherent challenge in distribution networks as opposed to transmission networks. The *State Estimation Algorithm* functional block then seeks to determine a unique solution or *system state*. Also, critical to the overall state estimation functionality and final determination of the system state is the *Bad Data Identification and Processing* functional block that uses statistical techniques (e.g., Chi-square Test) to identify and filter out "noise" which may be related to inaccuracies in measurement meters and/or communication system failures.

Finally, the *Human/Machine Interface* functional block relates to the software and hardware utilized to visualize and otherwise monitor and control the power system.

Further challenges beyond lack of metering, are those associated with *topology errors* and *false data injection attacks*. The terms and consequences of *lack of observability*, *topology errors* and *false data injection attacks* will be explained in later sections of this paper.

Figure 3 summarizes the key characteristics of the *"conventional" state estimator* based upon *weighted least squares.*

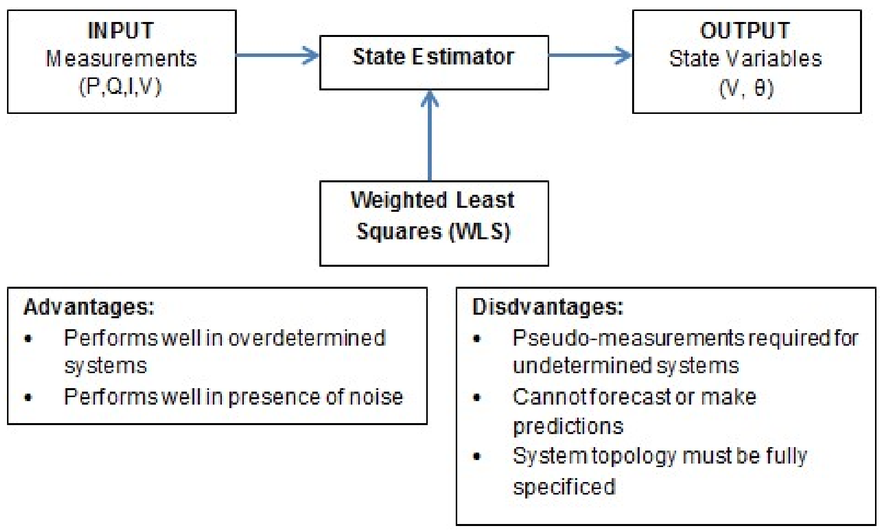

Figure 3 - State Estimator Overview

Note that the *INPUT* are typically measurements of P (Real Power), Q (Reactive Power), I (Current Flows), Voltage Magnitudes, and the *OUTPUT* state variables are typically voltage magnitudes and voltage phase angles at all buses. With these two *state variables*, it is then possible to determine the remaining parameters such as Real and Reactive Power Injections and Current flow.

Note that one of the buses can be established as the *reference bus or slack bus*. Thus, if Bus 1 is established as the reference bus, then the phase angle for Bus 1 can be removed from the vector representation. Therefore, if there are *n* buses in the network, the total number of states is given as *2n – 1*.

It is important to note that conventional state estimation applies only to *overdetermined systems*. Overdetermined systems are those in which the number of measurements exceeds the number of states. This critical and limiting requirement for application of conventional state estimation can be summarized in the following criteria:

- If the number of measurements is m, and the number of states is 2n-1, then in state estimation, m > 2n -1
- If m = 2n -1, the problem reduces to a power flow solution

Thus, as stated previously, distribution systems with limited measurement devices are inherently not overdetermined systems. For such *underdetermined systems* that may be either transmission or distribution networks lacking sufficient metering, *observability* is reduced and as indicated in Table 3, the state estimation algorithm must rely upon *pseudo-measurements*.

C. State Estimation Applied to Smart Distribution Systems

The authors of [12] provide a survey on state estimation techniques and challenges in so-called "smart distribution systems". This survey summarizes most of the essential concepts considered to this point regarding the following topics: Conventional mathematical formulation based upon an iterative algorithm utilizing weighted least squares or similar estimator, Application of pseudo-measurements to mitigate lack of sufficient metering to enable system observability, Consideration of optimal meter placement given the relatively limited metering, Network topology issues and effects, Impacts of renewable penetration, and Cybersecurity concerns. The paper goes further to make a distinction between "conventional" state estimation that is considered *analytical* and *deterministic* and "modern" state estimation that is considered *data driven* and *probabilistic*.

Regarding conventional state estimation, various "robust state estimators" are presented along with their pros and cons. For example, Generalized Maximum-likelihood (GM) has pros of Robust against bad data and cons of Parameter selection sensitivity.

Two major categories of data driven approaches are identified as alternatives to conventional state estimation based upon the previous list of estimators: *Probabilistic and Statistical Approaches* that employ spatial/temporal correlation and historical probability distributions, used widely for pseudo-measurement generation and uncertainty assessment, and *Learning-Based Approaches I Machine learning algorithms* that address problem of active/reactive power *pseudo-measurement generation and uncertainty assessment.*

Related to the recommendations of notable research directions, the paper [13] presents previous work in the area of state estimation for real-time monitoring of distribution systems. While the work presented is based upon weighted least squares estimation, it shows the close correlation of state estimation accuracy to the initial starting point selected and accuracy of the forecasted loads.

Thus, an important takeaway from the work presented in [12] and [13] collectively is the idea of establishing a *hybrid process* involving classical state estimation algorithms and data-driven forecasting.

The data-driven portions would support the classical state estimation algorithm by providing a better starting point than a typical "flat start", higher probability of convergence, and more accurate pseudo-measurements than those queried from large historical data repositories.

The design of an off-line planning method to enable real-time monitoring and control in systems with limited observability is considered in [14] through consideration of robust measurement placement for distribution system state estimation. This paper proposes a robust measurement placement model to maximize estimation accuracy for DSSE over a wide-range of worst case operating conditions.

The problem is formulated as a mixed-integer semi-definite programming problem (MISDP). The authors seek to avoid combinatorial complexity through a convex relaxation, followed by a local optimization method. The approach demonstrates that accuracy of DSSE can be enhanced significantly by placing a limited number of measurements in *optimal locations*. Again, the approach taken, can be considered a hybrid approach of classical state estimation with updated probabilistic and statistical components that seek to minimize the effect of lack of observability on the weighted least squares estimator.

The paper presented in [15], provides a linear state estimation formulation for smart distribution systems. The authors assume the availability of synchro-phasors which yield direct voltage phasors at bus locations. Line power flows and current magnitudes are then able to be ascertained via the direct quantities available. The authors show that availability of direct voltage phasors effectively *linearizes the h(x) coefficient matrix* used in classical state estimation so that the result is a linear, non-iterative state estimation solution. Results confirm low computational burden, accommodation of meshed networks and avoidance of convergence issues which may occur in dealing with practical distribution systems with high r/x ratios. It should be noted, however that to achieve the results, the following must be maintained by the synchro phasors:

- Resolution Requirement: +/- 1 µS which corresponds to 0.0216 degree phase error in a 60 Hz system
- Maximum Allowable Total Vector Error (TVE) : 1.0% when maximum phase error is 0.57 degrees

The authors in [16] present a branch-estimation-based state estimation method for radial distribution systems. While this approach utilizes many of the conventional or classical state estimation techniques, it has the ability to handle most kinds of *real-time measurements* by decomposing the weighted least squares problem into a series of weighted least squares problems such that each sub-problem deals with single-branch estimation. The establishment of "zones" is an idea, where the entire distribution system can be comprised of much simpler single-branches and each zone will then correspond to a weighted least squares sub problem. Ref [16] proposes two main parts: *load allocation and state estimation*. The load allocation portion is considered to be a real-time load modeling technique that incorporates use of customer class curves and provides a measure of the uncertainty (statistics) in the estimates. The purpose of this portion is to produce pseudo-measurements with a higher level of accuracy in real-time than historical data that must be retrieved from a large data repository. The state estimation portion then utilizes the pseudo-measurements that ensure observability and follows a traditional weighted least squares technique that is applied to each "zone".

The authors propose that a forward/backward sweep scheme based upon this method would allow state estimation to be performed accurately for large-scale practical distribution systems while not requiring sparse matrix techniques.

D. *Challenges of Applying Conventional State Estimation Utilizing Weighted Least Squares to Distribution Systems*

The most common conventional state estimation algorithm is based upon the *Weighted Least Squares (WLS)* algorithm.

The following list provides some of the characteristics of distribution systems that pose major challenges to the direct application of conventional state estimation based upon weighted least squares:

- Radial Topology with bi-directional power flow
- Lack of adequate quality and quantity of measurement devices resulting in underdetermined systems and thus reduced observability
- Unbalanced Lines and Loads resulting in the need to consider all phases in the state estimator algorithm
- Unpredictability of energy sources injecting power back onto the grid (i.e. intermittent sunlight and wind, electric vehicles, etc.).
- Variability in the timing of power utilization throughout the day
- Low X/R ratios which do not allow for neglecting resistances
- Substantial number of nodes, combined with the need to consider all phases, result in the need for acquisition, storage and processing of substantial amounts of data
- Excessive noises resulting from the variety and lack of standardization of communication schemes between metering devices and the central control stations

It should be noted that the limitations listed above are considered *"normal conditions"* inherent in all distribution systems. The addition of *"adverse conditions"* noted previously further strengthens the case for needed research of methods such as artificial neural networks to maintain data integrity of distribution system state estimation and thus the overall resiliency of the modern power grid.

E. *Lack of Observability in Distribution Systems*

In the context of this paper, *lack of observability* will be directly related to the inability to accurately measure and store system values (power, voltage magnitude, voltage phase angles and current flow) of a distribution system due to lack of measurement devices, failures in devices, communication failures and/or malicious attacks that would also fall into the category of False Date Injection Attacks.

While there are increasing advances in and application of Phasor Management Units (PMUs) and so-called "smart-meters", in this paper, there will not be an assumption that these devices are available at every bus location of a practical distribution system.

Thus, distribution system state estimation will be considered to be fundamentally challenged by *lack of observability*.

F. *Topology Errors in Distribution Systems*

In the context of this paper, *topology errors* will be directly related to errors in determination of system state values due to inaccurate determination of system breaker position. More generally, these errors could relate to incorrect determination of any device that involves switching or tap positioning.

The false status of system breakers could result from failures in devices, communication failures and/or malicious attacks that would also fall into the category of False Date Injection Attacks.

Thus, distribution system state estimation will also be considered to be fundamentally challenged by *topology errors*.

G. False Data Injection Attacks in Distribution Systems

In the context of this paper, *false data injection attacks* will refer to malicious attempts to alter data within distribution systems such that the *true system state* is made inaccurate. The goal of such attacks could be financial, such as controlling aspects of the power market or sabotage to the security of the power system resulting in power outages.

It should be noted that with advances in smart grid metering and reliance on digital communications, the susceptibility of the power grid to false data injection attacks will continue to be a growing security concern.

Thus, distribution system state estimation will also be considered to be fundamentally challenged (even threatened), by *false data injection attacks*.

H. Conventional Feed-Forward Multi-Layer Perceptron Networks (MLPs)

This type of network is considered the *conventional* or *classical* neural network model. Figure 4 shows a "perceptron", the fundamental building block of neural networks.

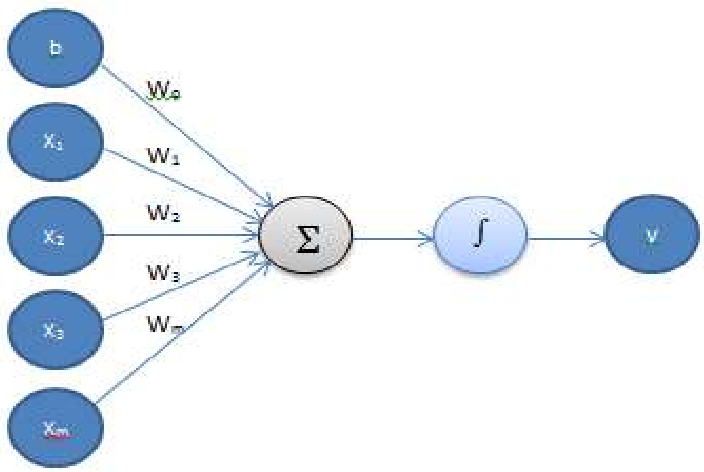

Figure 4 - Perceptron Building Block of MLP Networks

Figure 5 depicts the functional blocks of a MLP network model.

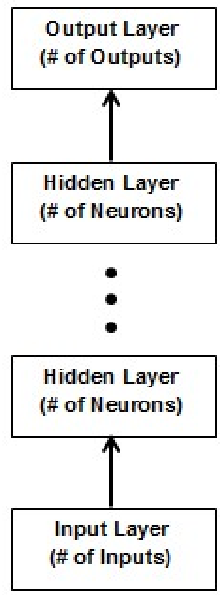

Figure 5 - Multi-Layer (MLP) Model Functional Representation

This type of network is considered a reasonable model for *regression* and *classification* problems. However, it has limited ability to predict or forecast sequence or time-series data as it does not maintain and share features between layers.

This type of neural network is also limited to how “deep” they can be in terms of number of layers that would otherwise enable them to solve more complex problems with greater accuracy.

Even with the noted limitations, this network type has promises to overcome many of the limitations of weighted least squares based state estimation. The principal advantage of this network type is the promise to accurately learn the mapping of inputs to outputs for a regression problem without the requirement of complex and large number of equations that would be necessary to perform non-linear regression on large distribution systems.

I. Convolutional Neural Networks (CNNs)

This type of network is considered to be an improvement upon the classical MLP architecture in that it *learns directly from the input data* and thus does not require a target dataset during training. Figure 6 shows the general structure for a CNN model.

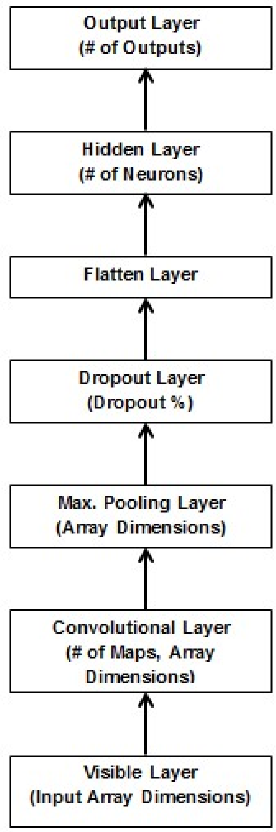

Figure 6 - Convolutional Neural Network (CNN) Model Functional Representation

The fundamental layer types are as follows:

- Convolutional Layers
  - Comprised of Filters and Feature Maps
  - Filters correspond to neurons of the layer
  - Filters have weighted inputs and produce outputs like a neuron
  - Filters input size is fixed and is a “window” for convolution
  - Feature Maps contain current values within the moving filter window
- Pooling Layers
  - Down-sample and consolidate features learned from previous feature maps
  - Serve to generalize or compress features selected
  - Reduce overfitting of model training
  - Simple functionality – selection of either maximum or average of input value to establish a new compressed feature map
- Dropout Layers
  - Used between other layers to further reduced overfitting not completely eliminated by pooling layers by randomly excluding neurons
  - Specified by a Dropout Percentage
- Flatten Layers
  - Converts multidimensional arrays to vectors that can be sent to fully connected layers for final processing by activation functions
- Fully Connected Layers
  - Normal flat feedforward neural network layer
  - Contain a ‘softmax’ or nonlinear activation function to output probabilities of predicted classes
  - Utilized at the end of network to create combinations of nonlinear features used for predictions

While primarily used in image/object detection and classification, computer vision and natural language processing, the research surrounding this paper will investigate the feasibility of this network type to perform *regression* so as to detect and correct data errors imposed upon distribution state estimation. Regression in this context is the determination of numerical values such as the predicted system state values or the numerical values indicating the error and/or variance between actual and predicted values.

The principal advantage of this network type is its ability to automatically learn and generalize features from the input data.

J. Recurrent Neural Networks (RNNs)

This type of network is also considered to be an improvement upon the classical MLP architecture in that it maintains an internal state (memory). There are three primary variants of RNNs:

- Bidirectional Recurrent Neural Networks (BRNN):
  - RNNs that utilize future data along with data from previous inputs to improve accuracy
- Long Short Term Memory Networks (LSTM):
  - Discussed in more detail in the next section.
- Gated Recurrent Units (GRUs):
  - Like LSTMs, overcome short-term memory limitations of the basic RNN model
  - Uses hidden states instead of "cell state" utilized by LSTMs
  - Contains *reset* and *update gates* to control what information is retained and how much of this information to use for making predictions.

The principal advantage of this network type is that it maintains and passes features between layers, and thus very deep structures can be developed without the negative effects of *exploding or vanishing gradients*.

K. Long Short-Term Memory Networks (LSTMs)

This network is a type of RNN that can *learn long-term dependencies* between time steps of input sequence data by "remembering" the state between predictions. The following operations provide more details on the internal architecture of the LSTM unit.

- Step 1 – "Forget Gate" – Determines and eliminates previous information deemed as irrelevant and thus not useful
- Step 2 – "Store Gate" – Determines what new information to maintain as new candidate values.
- Step 3 – "Update Gate" -- Updates old cell state to new cell state
  - Multiply old cell state by f_t , forgetting things that were decided to be forgotten earlier.
  - Then scale new candidate values by how much it was decided to update each state value
- Step 4 – "Output Gate" – Determines what is to be output for the next step
  - Output will be based up cell state, but will be a filtered version
    - First run a sigmoid layer to decide what parts of a cell state to output
    - Then put the cell state through *tanh* activation function to push values between -1 and 1 and multiply it by the output of the sigmoid gate so that only the desired parts are output

L. Hyper-Parameters Optimization

This research aims to provide an optimization method to determine the optimal hyper-parameters for desired performance metrics. Hyper-parameters include model parameters such as number of hidden layers and number neurons in a layer, and algorithm parameters such as adjustable learning rate. Hyper-parameters may be obtained using optimization methods such as grid search method, genetic algorithms, Bayesian optimization method, etc.

III. POWER FLOW SIMULATIONS

A. Selection of Base Distribution for Simulation

An IEEE 34 Bus Test Feeder radial distribution system was selected as the *base test distribution system*. It is shown in Figure 7 [17].

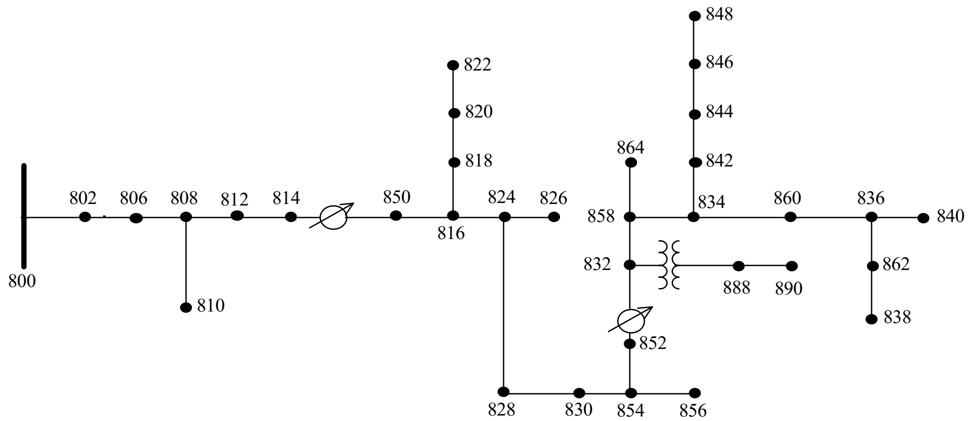

Figure 7 - IEEE 34 Node Test Base Distribution System [17]

B. Measurement Points and Quantities

For purposes of *training a supervised neural network*, it was decided that the power (real and reactive) at each bus for all 3 phases would be measured and deemed the *"input"* dataset. The voltage and phase angle at each bus for all 3 phases were selected to be measured and deemed the *"target"* dataset**.**

The selected measurement points and quantities are shown in Figure 8. The labels corresponding to the "Key" represent either a *power or voltage monitor,* which is similar to a physical meter and will be discussed in more detail later.

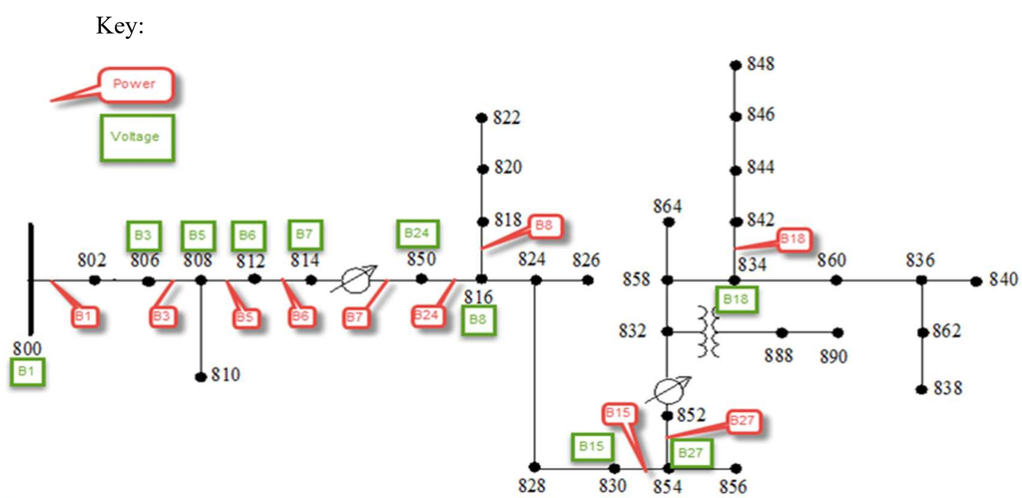

Figure 8 - IEEE 34 Node Test Base Distribution System Measurement Points

Table 1 and Table 2 provide a description of the monitors, their locations in the test distribution systems and the quantities they measure.

Note that power monitors capture the *real and reactive power flow* along the lines between specific nodes as indicated in Table 1. Likewise, voltage monitors capture the *voltage magnitude and voltage phase angle* at specific nodes as indicated in Table 2.

Table 1 - Power Monitor Descriptions and Locations

Monitor	**Line Element**	**From Node**	**To Node**	**Quantities**
B01_power	L1	800	802	Phase A,B,C P (kW) and Q(kVAR)
B03_power	L3	806	808	Phase A,B,C

				P (kW) and Q(kVAR)
B05_power	L5	808	812	Phase A,B,C P (kW) and Q(kVAR)
B06_power	L6	812	814	Phase A,B,C P (kW) and Q(kVAR)
B07_power	L7	814	850	Phase A,B,C P (kW) and Q(kVAR)
B24_power	L24	850	816	Phase A,B,C P (kW) and Q(kVAR)
B08_power	L8	816	818	Phase A P(kW) and Q(kVAR)
B15_power	L15	830	854	Phase A,B,C P (kW) and Q(kVAR)
B18_power	L18	834	842	Phase A,B,C P (kW) and Q(kVAR)
B27_power	L27	854	852	Phase A,B,C P (kW) and Q(kVAR)

Table 2 - Voltage Monitor Descriptions and Locations

Monitor	Line Element	Node	Quantities
B01_voltage	L1	800	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B03_voltage	L3	806	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B05_voltage	L5	808	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B06_voltage	L6	812	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B07_voltage	L7	814	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B24_voltage	L24	850	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B08_voltage	L8	816	Phase A to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B15_voltage	L15	830	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B18_voltage	L18	834	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)
B27_voltage	L27	854	Phase A,B,C to Neutral Voltage Mag. (V) and Phase Angle (degrees)

C. Power Flow Simulation to Establish Baseline Datasets

For purposes of performing a power flow simulation of the test feeder system to gather the power and voltage at each bus, *OpenDSS* from Electric Power Research Institute, Inc. (EPRI) was chosen. Note that the convention in OpenDSS is that Phase-1, Phase-2 and Phase-3 represent phases a, b, and c, respectively.

It was decided that the loads within the test distribution feeder would be varied over a time period of a year (8760 hours) to yield a time-series dataset corresponding to the power and voltage as discussed previously.

To vary the base loads in a realistic manner, historical data from the *Electric Reliability Council of Texas (ERCOT)* as obtained. The load data for the entire ERCOT grid for every hour of the entire year of 2018 was selected. The ERCOT load dataset was then used to realistically scale the power (P and Q values) at each node that contains a load to establish the needed variation over a period of a year.

Note that “ERCOT” will be used as the baseline load profile, and all references to ERCOT datasets have their origin from the baseline power flow simulation of the test distribution system performed with varying loads according to this load profile.

OpenDSS (version 9.1.0.1, 64-bit build) was then utilized to perform a power flow simulation of the test feeder distribution system with varying load, and the power and voltage monitors indicated in Table 1 and Table 2 exported the power and voltage datasets. This exported data would serve as the input and target datasets from the test system under normal conditions. Training and testing of the neural networks types would be based upon this data.

D. *Power Flow Simulation to Establish Previously Unseen Datasets*

The previous steps related to performing a power flow simulation with OpenDSS were repeated with a different load profile to establish *previously unseen data* for validating the various neural network types.

Note that "COAST" will be used in descriptions of datasets that have their origin from the power flow simulation of the test distribution system performed with varying loads according to this load profile.

IV. MULTILAYER PERCEPTRON MODEL (MLP)

A. *Network Model Parameters*

- Number of inputs in visible layer: 56 (Held constant for Trials 1 – 11)
  - Power Monitors:
    - B01_power: 6 features (P and Q values for 3 phases)
    - B03_power: 6 features (P and Q values for 3 phases)
    - B05_power: 6 features (P and Q values for 3 phases)
    - B06_power: 6 features (P and Q values for 3 phases)
    - B07_power: 6 features (P and Q values for 3 phases)
    - B24_power: 6 features (P and Q values for 3 phases)
    - B08_power: 2 features (P and Q values for 1 phase)
    - B15_power: 6 features (P and Q values for 3 phases)
    - B18_power: 6 features (P and Q values for 3 phases)
    - B27_power: 6 features (P and Q values for 3 phases)
  - Number of hidden layers and number of neurons per hidden layer:
    - Adjusted for each trial according to Table 3
- Number of output layer: 56 (Held constant for Trials 1 – 11)
  - Voltage Monitors:
    - B01_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B03_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B05_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B06_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B07_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B24_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B08_voltage: 2 features (Mag. and Phase values for 1 phase)
    - B15_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B18_voltage: 6 features (Mag. and Phase values for 3 phases)
    - B27_voltage: 6 features (Mag. and Phase values for 3 phases)

B. *Hyper-Parameters (Held constant for Trials 1-11)*

- Activation Function per Layer : Hyperbolic Tangent (tanh)
- Loss Function: Mean Squared Error
- Optimizer: Stochastic Gradient Descent
- Batch Size = 10

V. TRIAL RESULTS

The following results are for MLP topologies trained and tested on ERCOT data and validated on COAST data.

A. *Trails 1 - 11*

Table 3 presents training, testing and validation root-mean squared errors for nine MLP model architectures. As indicated in Table 3, the number of hidden layers and number of hidden layer neurons were varied. The number of input and output layer neurons was held constant at 56 neurons to correspond to the number of input and output features.

As indicated in this table, 70% of the ERCOT data was used for training and 30% was held out for testing. The "COAST Act. vs. Est" column shows results for the various architectures of the MLP when predicting output voltages and phase angles for COAST data that has never been seen by the neural network.

Table 3 - Trial Results for Baseline MLP Model without Hyper-Parameter Optimization

Trial	Input Layer	Hidden Layers	Output Layer	Train RMSE (70%)	Test RMSE (30%)	COAST Act. vs. Est. RMSE
1	56 Neurons	1 Layer 56 Neurons	56 Neurons	0.140927	0.142162	0.323075
2	56 Neurons	1 Layer 112 Neurons	56 Neurons	0.140486	0.136711	0.323293
3	56 Neurons	1 Layer 224 Neurons	56 Neurons	0.137222	0.137328	0.322124
4	56 Neurons	10 Layers 56 Neurons	56 Neurons	0.092110	0.092036	0318610
5	56 Neurons	10 Layers 112 Neurons	56 Neurons	0.091231	0.090394	0.317231
6	56 Neurons	10 Layers 224 Neurons	56 Neurons	0.089581	0.089328	0.333226
7	56 Neurons	20 Layers 56 Neurons	56 Neurons	0.085845	0.087380	0.309943
8	56 Neurons	20 Layers 112 Neurons	56 Neurons	0.082363	0.082748	0.314640
9	56 Neurons	20 Layers 224 Neurons	56 Neurons	0.077807	0.078516	0.314956
10	56 Neurons	50 Layers 224 Neurons	56 Neurons	0.076469	0.078396	0.313638
11	56 Neurons	100 Layers 224 Neurons	56 Neurons	0.079144	0.079668	0.307243

B. Error Distributions for Baseline MLP Network

- Trial 1 - 1 Layer 56 Hidden Neurons:

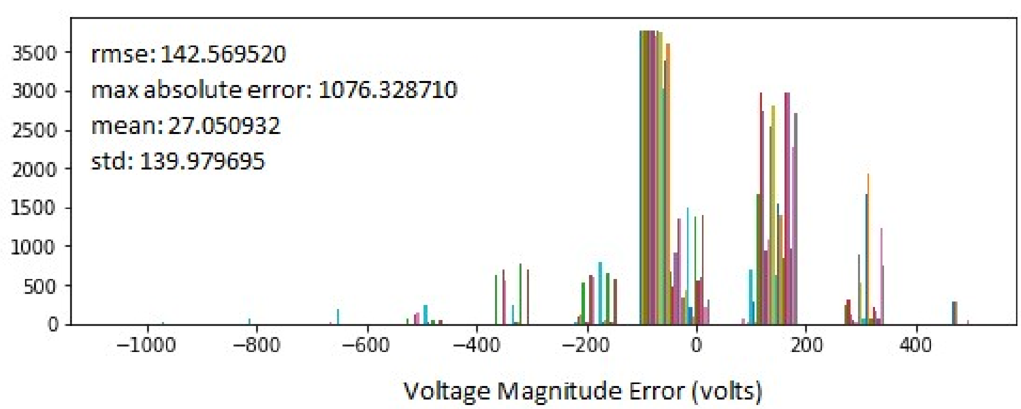

Figure 9 - Voltage Magnitude Error Distribution for MLP without Hyper-Parameter Optimization

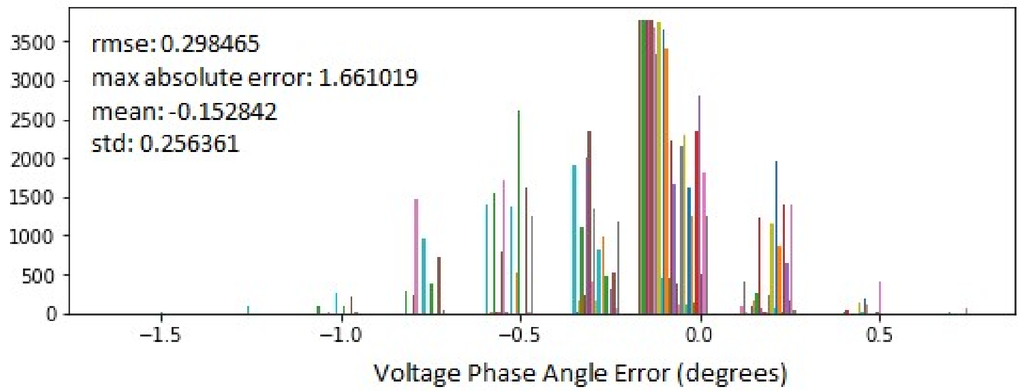

Figure 10 - Voltage Phase Angle Error Distribution for MLP without Hyper-Parameter Optimization

- Trial 9 - 20 Layers 224 Hidden Neurons:

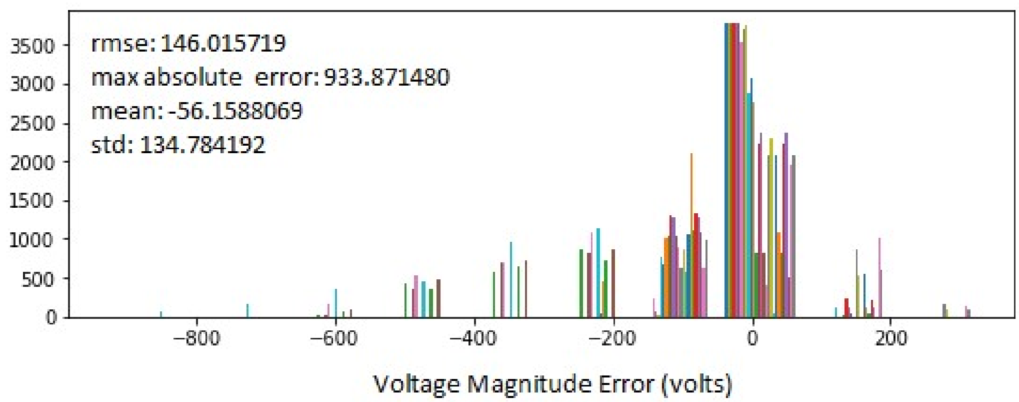

Figure 11 - Voltage Magnitude Error Distribution for MLP without Hyper-Parameter Optimization

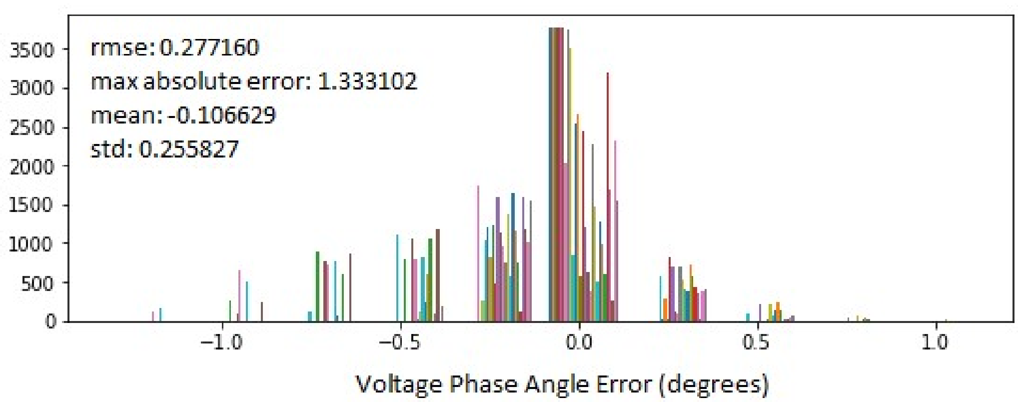

Figure 12 - Voltage Phase Angle Error Distribution for MLP without Hyper-Parameter Optimization

- Trial 10 - 50 Layers 224 Hidden Neurons:

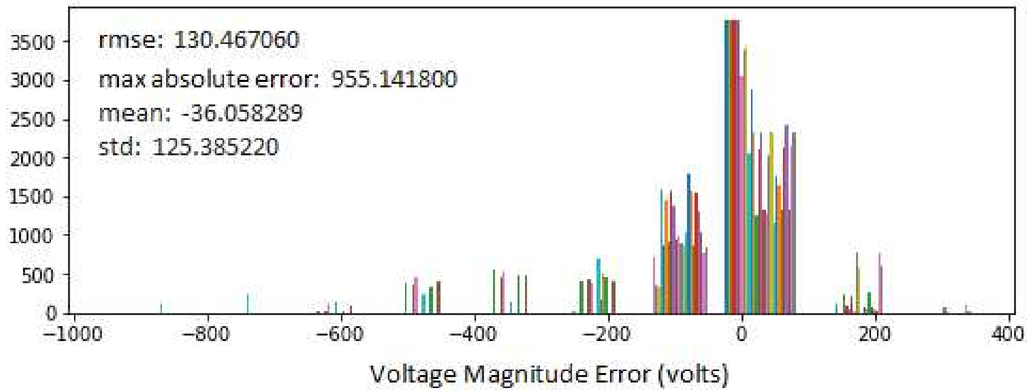

Figure 13 - Voltage Magnitude Error Distribution for MLP without Hyper-Parameter Optimization

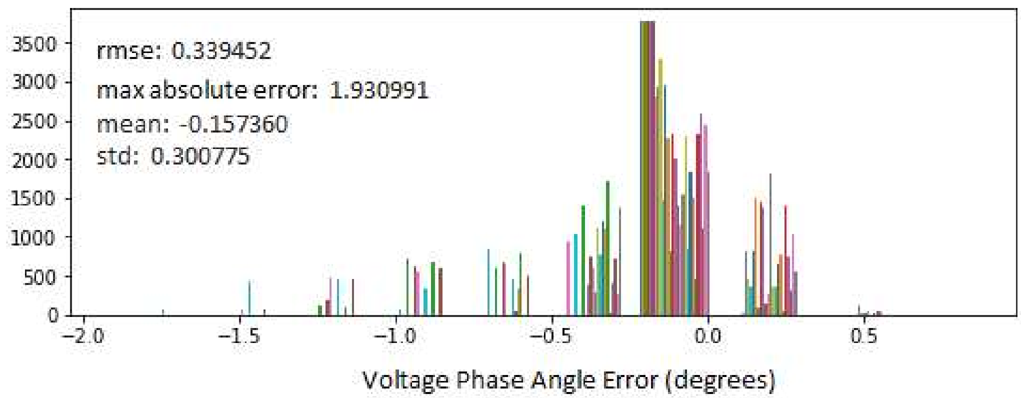

Figure 14 - Voltage Phase Angle Error Distribution for MLP without Hyper-Parameter Optimization

- Trial 11 - 100 Layers 224 Hidden Neurons:

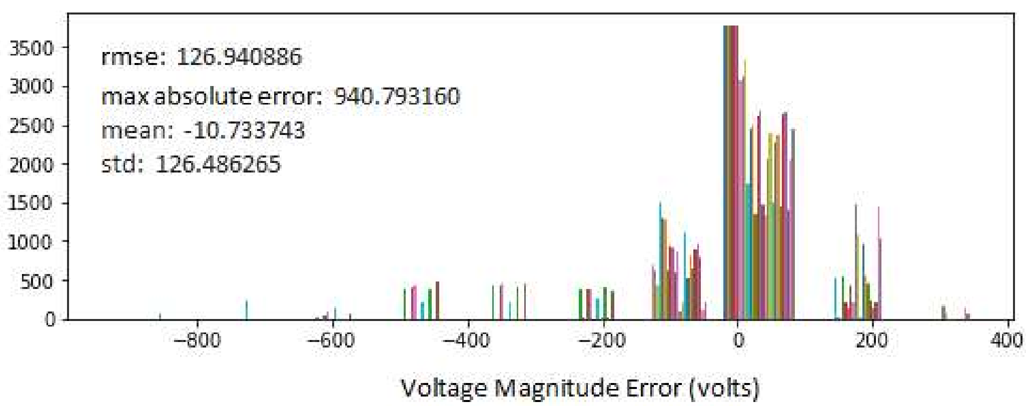

Figure 15 - Voltage Magnitude Error Distribution for MLP without Hyper-Parameter Optimization

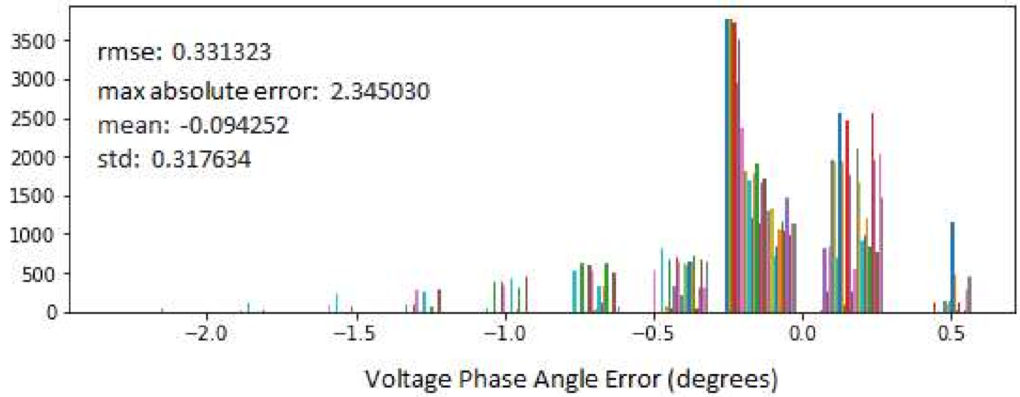

Figure 16 - Voltage Phase Angle Error Distribution for MLP without Hyper-Parameter Optimization

The results presented in Table 3 show that with adjustment of the number of hidden layers and number of hidden layer neurons, a feedforward multilayer perceptron model (MLP) shows promise in terms of serving as a fully data-driven distribution system state estimator.

Figures 9-16 reveal that the error distributions of an un-optimized MLP can be modeled as approximately Gaussian or more accurately as mixed Gaussian.

VI. Conclusions

State estimation applied to electric power systems has been proposed since the early 1970s. The application of state estimation was primarily made to transmission systems as opposed to distribution systems. Classical or conventional state estimation was based upon an iterative algorithm to minimize error utilizing estimators such as weighted least squares. There are challenges to develop state estimation algorithms for power distribution systems due to inherent system unbalance among phases, bi-directional power flow and more recently, and dynamics and uncertainty associated with distributed energy resources (i.e. photovoltaic ,wind, and electric vehicles).

This research focuses on the data-driven approaches to the state estimation problem that employ the application of machine learning and neural networks in general and deep learning models in particular to mitigate the challenges associated with the direct application of conventional analytical approaches. Initial results based on MLPs are presented. The state estimation problem was staged with a power flow simulation of an IEEE 34 Node Test Feeder. This simulation provided *input data* consisting of real and reactive power flows between nodes and *output or target data* consisting of voltage magnitudes and phase angles at nodes for use in training MLPs.

In future research, we will also examine CNN and LSTM architectures and hybrid models that may contain elements of conventional state estimation methods and various combinations of MLP, CNN and LSTM architectures. After training using the previously gathered input and output datasets from the power flow simulation, each of these model types will be evaluated in terms of their ability to perform regression in order to predict voltage magnitudes and phase angles as outputs given previously unseen real and reactive power as inputs.

Initially, the models will not be optimized and their configuration will follow an ad-hoc or heuristic approach. Future research will evaluate the ability to optimize so-called *hyper-parameters* of each model type to determine a methodical approach to model configuration. Consideration of application to larger distribution networks will also be made.

References

[1]. Schweppe, Fred C, and Wildes, J. "Power System Static-State Estimation, Part I: Exact Model," *IEEE Transactions on Power Apparatus and Systems* PAS-89.1 (1970): 120-25.

[2]. Schweppe, Fred C, and Rom, Douglas B. "Power System Static-State Estimation, Part II: approximate Model," *IEEE Transactions on Power Apparatus and Systems* PAS-89.1 (1970): 125-30.

[3]. W. Fan et al., "A CVR On/Off Status Detection Algorithm for Measurement and Verification," 2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 2021, pp. 1-5.
[4]. Wen Fan, Yuan Liao, “Microgrid operation optimization considering storage devices, electricity transactions and reserve,” International Journal of Emerging Electric Power Systems, vol. 20, No. 5, September 2019.
[5]. Wen Fan, 2019, Advanced Fault Area Identification and Fault Location for Transmission and Distribution Systems, Dissertation, USA.
[6]. Wen Fan and Yuan Liao, “Fault identification and location for distribution network with distributed generations,” International Journal of Emerging Electric Power Systems, vol. 19, no. 3, May 2018.
[7]. Schweppe, Fred C. "Power System Static-State Estimation, Part III: Implementation." *IEEE Transactions on Power Apparatus and Systems* PAS-89.1 (1970): 130-35.
[8]. Shivakumar, N.R, and Jain, A. "A Review of Power System Dynamic State Estimation Techniques," 2008 Joint International Conference on Power System Technology and IEEE Power India Conference (2008): 1-6.
[9]. Krumpholz, G. R, Clements, K. A, and Davis, P. W. "Power System Observability: A Practical Algorithm Using Network Topology," *IEEE Transactions on Power Apparatus and Systems* PAS-99.4 (1980): 1534-542.
[10]. Urban Kuhar, Gregor Kosec, and Ales Svigelj, *Observability in Distribution Systems*, January 2020 – Full Textbook.
[11]. Electric Power Research Institute (EPRI), *Electrical Power System Resiliency: Challenges and Opportunities*, EPRI White Papers, February 2016.
[12]. Kaveh Dehghanpour, Zhaoyu Wang, Jianhui Wang ,Yuxuan Yuan and Fankun Bu, “A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems,” *IEEE Transactions on Smart Grid*, Vol. 10, No. 2, March 2019.
[13]. Baran, M.E, and Kelley, A.W. "State Estimation for Real-time Monitoring of Distribution Systems," *IEEE Transactions on Power Systems* 9.3 (1994): 1601-609.
[14]. Yao, Yiyun, Liu, Xuan, and Li, Zuyi. "Robust Measurement Placement for Distribution System State Estimation." *IEEE Transactions on Sustainable Energy* 10.1 (2019): 364-74. Web.
[15]. Haughton, Daniel A, and Heydt, G. T. "A Linear State Estimation Formulation for Smart Distribution Systems," *IEEE Transactions on Power Systems* 28.2 (2013): 1187-195.
[16]. Youman Deng, Ying He, and Boming Zhang. "A Branch-estimation-based State Estimation Method for Radial Distribution Systems," *IEEE Transactions on Power Delivery* 17.4 (2002): 1057-062.
[17]. IEEE Test Feeders: https://site.ieee.org/pes-testfeeders/resources/.