

DATA-DRIVEN APPROACHES FOR DISTRIBUTION GRID MODERNIZATION:
EXPLORING STATE ESTIMATION, PSEUDO-MEASUREMENT GENERATION AND
FALSE DATA DETECTION

by

Sepideh Radhoush

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

December 2023

©COPYRIGHT

by

Sepideh Radhoush

2023

All Rights Reserved

DEDICATION

To my Father, Mother and Brother who have defined the true meaning of Unconditional Love and Support to me throughout my life.

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor Dr. Bradley M. Whitaker for his constant support from day one I joined his team who helped me with all obstacles, confusions and ups and downs. Our meetings were always very helpful and the resources he shared with me helped me a lot to reach a solution whenever I was puzzled. This research would never have been successful without his constant support.

I am deeply grateful to my respected co-advisor, Professor Hashem Nehrir, whose invaluable guidance, knowledge, and unwavering support, even in retirement, have been instrumental in the completion of my dissertation. His mentorship and the generous support of his NSF funds have truly enriched my academic journey.

I am also grateful to the ECE program, and Professor Todd J. Kaiser for their support during my Ph.D. journey and also very grateful to Dr. Hongwei Gao, and Dr. Ying Zhang for their kindness to be on my committee and giving me constructive comments and directions to enrich my research.

I am also thankful to my friends, Trevor Vannoy and Kaveen Liyanage, for their help in developing and performing my idea during my Ph.D. degree.

TABLE OF CONTENTS

1. CHAPTER ONE	1
INTRODUCTION	1
2. CHAPTER TWO	4
LITERATURE REVIEW	4
Machine Learning (ML) Approach	9
Supervised Machine Learning Method	11
Assessment of Regression and Classification Method	20
Machine Learning Approach for Generating Pseudo-Measurements	22
Machine Learning Approach for Hybridization Traditional and Smart Measurements	24
Machine Learning Approach for Distribution System State Estimation	26
Machine Learning Approach for False Data Injection Attacks in a Distribution Grid	31
Conclusion	34
3. CHAPTER THREE	36
INCORPORATING TEMPORAL AND METEOROLOGICAL DATA FOR GENERATING PSEUDO-MEASUREMENTS IN ACTIVE DISTRIBUTION POWER NETWORKS	36
Methodology	38
Case Study	43
Conclusion	48
4. CHAPTER FOUR.....	50
DISTRIBUTION SYSTEM STATE ESTIMATION AND FALSE DATA INJECTION ATTACK DETECTION WITH A MULTI-OUTPUT DEEP NEURAL NETWORK.....	50
Power System State Estimation	53
WLS formulation for State Estimation	53
False Data Injection Attacks (FDIAs).....	56
Machine Learning Approach to Detect FDIAs	57
Methodology	59
Results.....	63
Case Study I	63
Case Study II.....	74
Conclusion	83
5. CHAPTER FIVE	85
DISTRIBUTION SYSTEM STATE ESTIMATION USING HYBRID TRADITIONAL AND ADVANCED MEASUREMENT FOR GRID MODIRNIZATION	85
Traditional and Smart Measurements	86
Distribution System State Estimation	88
Methodology	91
Results.....	95
Case Study I	95
Case Study II.....	98

TABLE OF CONTENTS CONTINUED

Conclusion	102
6. CHAPTER SIX.....	105
CONCLUSION.....	105
Publication	109
REFERENCES CITED.....	110

LIST OF TABLES

Table	Page
1. Table 3- 1. .Input and output of the proposed method	44
2. Table 3- 2. Hyperparameters of each base learner for the proposed and random split methods.....	44
3. Table 3- 3. RMSE (%) values for voltage magnitude estimation for the proposed method and random split.....	48
4. Table 4- 1. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification- only DNN model when $N_a = 1200$ (10%), and $z_a = (1 \pm 0.05)z, (1 \pm 0.1)z$	67
5. Table 4- 2. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification-only DNN model when $N_a = 600$ (5%), $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$	70
6. Table 4- 3. . Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification-only DNN model when $N_a = 1200$ (10%), $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$	72
7. Table 4- 4. Execution time for the proposed, regression-only, and WLS methods.	74
8. Table 4- 5. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification- only DNN model when $N_a = 876$, and $z_a = (1 \pm 0.05)z, (1 \pm 0.1)z$	77
9. Table 4- 6. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification-only DNN model when $N_a = 438$ (5%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$	79
10. Table 4- 7. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification- only DNN model when $N_a = 876$ (10%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$	81

LIST OF TABLES CONTINUED

11. Table 4- 8. Execution time for the proposed, regression-only, and WLS methods.....	83
12. Table 5- 1 DSSE results of voltage magnitudes and phase angles after grid modernization.	98
13. Table 5- 2. DSSE Results of Voltage Magnitudes and Phase Angles After Grid Modernization when the training model is updated every 10- or 20- time steps (twice or once per day).	102

LIST OF FIGURES

Figure	Page
14. Figure 2- 1. Smart power grid architecture.	5
15. Figure 2- 2. Visual representation of four ML categories: Supervised, Unsupervised, Semi-Supervised, and Reinforcement.	10
16. Figure 2- 3. Different SML techniques, categorized by regression and classification.	12
17. Figure 2- 4. Linear and polynomial (2 nd order) regression models for 1-dimensional inputs.	15
18. Figure 2- 5. Logistic Regression Model.	15
19. Figure 2- 6. (a) Support vector Regressor; (b) Support Vector Classification.	17
20. Figure 2- 7. Decision tree algorithm	18
21. Figure 2- 8. (a) A single-layer perceptron; (b) A multi-layer perceptron (DNN)	19
22. Figure 3- 1. Diagram of the proposed method.	39
23. Figure 3- 2. Training and testing diagram of the proposed method.	39
24. Figure 3- 3. (a) Month of a year, (b) Day of a week, and (c) Hour of a day.	40
25. Figure 3- 4. PV outputs over two consecutive days.	41
26. Figure 3- 5. A random forest model.	42
27. Figure 3- 6. IEEE Standard 69 bus Distribution System.	43
28. Figure 3- 7. MPE and RMSPE values for (a) the proposed method and (b) and random split method with no DG and four DGs.	46
29. Figure 3- 8. Active power injection at 08/29 at 8:00 AM (Pseudo-Measurements Generated).	46
30. Figure 3- 9. Active power injection at 11/04 at 1:00 PM (Pseudo-Measurements Generated).	47

LIST OF FIGURES CONTINUED

31. Figure 3- 10. Actual and predicted values for Bus 38.....	47
32. Figure 4- 1. General architecture of FDIAs on measurements and DSEE procedure.....	57
33. Figure 4- 2. DNN configuration.....	58
34. Figure 4- 3. (a) A regression DNN architecture for DSSE calculation; (b) A binary classification DNN model for FDIAs detection on measurements; (c) the proposed DNN model configuration to perform DSSE calculation and FDIAs on measurements simultaneously.....	60
35. Figure 4- 4. IEEE standard 33 bus distribution network.	63
36. Figure 4- 5. Original and attacked measurements for 120 samples when $z_a = (1 \pm 0.05)z, (1 \pm 0.1)z$	65
37. Figure 4- 6. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 1200$ (10%), and $z_a = (1 \pm 0.05)z, (1 \pm 0.1)z$	68
38. Figure 4- 7. Original and attacked measurements for 103 samples when $z_a = \sim \mathcal{N}(\mu z, \sigma z)$	69
39. Figure 4- 8. DSSE Results Obtained Using the Proposed Method, the Regression- only DNN Model, and WLS with or without Bad Data when $N_a = 600$ (5%), and $z_a = \sim \mathcal{N}(\mu z, \sigma z)$	71
40. Figure 4- 9. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 1200$ (10%), and $z_a = \sim \mathcal{N}(\mu z, \sigma z)$	73
41. Figure 4- 10. The modified IEEE Standard 69 Bus Distribution System.	75
42. Figure 4- 11. DSSE results obtained using the proposed method, the Regression- only DNN Model, and WLS with or without Bad Data when $N_a = 876$, and $z_a = (1 \pm 0.05)z, (1 \pm 0.1)z$	78
43. Figure 4- 12. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 438$ (5%), and $z_a = \sim \mathcal{N}(\mu z, \sigma z)$	80

LIST OF FIGURES CONTINUED

44. Figure 4- 13. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 876$, and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$	82
45. Figure 5- 1. (a) SCADA and PMU measurements in a power network, (b) SCADA and PMU measurements refresh rates.	88
46. Figure 5- 2. A Multi-output regression model to predict SCADA measurements. For the figure, it is assumed that the PMUs are installed in the DN at time step N_1 , so the procedure is performed through time step N_1-5	93
47. Figure 5- 3. DSSE calculations and updating a training model after PMU installation.....	94
48. Figure 5- 4. IEEE standard 33 bus distribution system before and after grid modernization. Prior to modernization, there is a PV on bus 8 and SCADA units on buses 3, 8, and 21. After modernization, an additional PV is installed on bus 29 and PMU devices are installed on buses 1, 16, and 31.....	95
49. Figure 5- 5. IEEE standard 69 bus distribution system before and after grid modernization. Prior to modernization, there is a PV on bus 20 and SCADA units on buses 33, and 53. After modernization, two additional PVs are installed on bus 38 and 64, and PMU devices are installed on buses 1, and 22.....	99

ABSTRACT

Distribution networks must be regularly updated to enhance their performance and meet customer electricity requirements. Advanced technologies and infrastructure—including two-way communication, smart measuring devices, distributed generations in various forms, electric vehicles, variable loads, etc.—have been added to improve the overall efficiency of distribution networks. Corresponding to these new features and structures, the continuous control and monitoring of distribution networks should be intensified to keep track of any modifications to the distribution network performance. Distribution system state estimation has been introduced for real-time monitoring of distribution networks. State estimation calculations are highly dependent on measurement data which are collected from measurement devices in distribution networks. However, the installation of measurement devices is not possible at all buses to ensure the distribution network is fully observable. To address the lack of real measurements, pseudo-measurements are produced from historical load and generation data. Available measurements, along with physical distribution network topology, are fed into a state estimation algorithm to determine system state variables. Then, state estimation results are sent to a control center for further processing to enhance distribution network operation. However, the accuracy of state estimation results could be degraded by false data injection attacks on measurement data. If these attacks are not detected, distribution network operation could be significantly influenced. Different methods have been developed to enhance a distribution network operation and management. Machine learning approaches have also been identified to be beneficial in solving different types of problems in a power grid. In this dissertation, machine learning is applied to three areas of distribution systems: generating pseudo-measurements, performing distribution system state estimation calculations, and detecting false data injection attacks on measurement data. In addition to addressing these areas individually, machine learning is used to simultaneously perform distribution system state estimation calculation and false data injection attack detection. This is done by taking advantage of conventional and smart measurement data at different time scales. The results reveal that the operation and performance of a distribution network are improved using machine learning algorithms, leading to more effective power grid modernization.

CHAPTER ONE

INTRODUCTION

New challenges and issues have been raised about Distribution Networks (DNs) while advanced technologies are integrated to DNs. These new features—including two-way communications, Distributed Generations (DGs), Electric Vehicles (EVs), smart measurement devices, etc.—convert a passive DN into an Active Distribution Network (ADN). DN monitoring is becoming important for understanding the performance of ADNs in the presence of these advanced tools. State estimation (SE) approaches have been introduced to estimate the current operating state of the DN. SE is a critical tool which makes a link between measurement devices and control centers for monitoring and control of all the electrical quantities in ADNs. SE techniques were developed for Transmission Networks (TNs) several years ago, but these techniques cannot be applied directly to DNs due to the different features between TNs and DNs. Such features include a high resistance-reactance ratio (r/x), a lack of sufficient measurement devices, radial or weakly meshed structure, and three-phase unbalanced operation. Therefore, appropriate SE techniques should be realized according to the structure and need of DNs.

In recent years, DNs are being upgraded gradually due to the limited installation of smart metering devices to obtain precise assessment and monitoring, but a new challenge has been reported in terms of combining traditional and smart measurements. Traditional and smart measurement data are available at different sampling rates; both information should be utilized in performing SE calculations to include their positive advantages. Some hybrid SE methods have been employed to combine traditional and smart measurement data in DNs.

The measurement data should satisfy two conditions: (1) enough measurement data must be collected and (2) the measurement information should be precise enough to obtain Distribution System State Estimation (DSSE) results.

DSSE calculations are directly dependent on the available measurements from the measurement tools over the DN. However, installation of measurement devices is not economically feasible at all nodes in the DN. So, generating pseudo-measurements from a limited available measurement and historical data has been introduced as a new tool to address the lack of measurements and improve the observability of DNs. Some conventional methods were developed for generating pseudo-measurements, but these methods are not as effective for ADNs since the dynamic behaviors of DNs are intensified in the presence of DGs and these dynamic behaviors cannot be addressed properly using traditional methods.

As mentioned earlier, the accuracy of the DSSE results is also strongly dependent on the measurement accuracy. However, the integrity of measurement information might be easily degraded by False Data Injection Attacks (FDIAs) on measurement information. FDIAs are a kind of cyber-attack that is introduced in a DN to falsify measurement data and disrupt a DN's normal operation. The accuracy of DSSE results and the quality of decision-making, which is processed in a control center might be diminished significantly if FDIAs cannot be well identified. FDIA detection methods have been introduced and developed to analyze measurement data and enhance the DN performance. The purpose of bad data detection (BDD) is to identify the faulty data in measurement information. However, since FDIAs can successfully bypass many traditional BDD methods, FDIAs cannot be commonly identified and errors can be introduced into the DSSE result.

In the last decade, different machine learning approaches have been developed and widely used in DNs for different purposes, such as generating pseudo-measurements, performing SE calculations, detecting FDIAs, analyzing reliability and resiliency, etc. Machine learning algorithms are based on the data obtained from the power networks, unlike model-based approaches which are based on network topologies as well as measurement data. The data-based approach of machine learning enables faster execution time and more accurate results than is possible using traditional methods in DNs.

In the rest of this dissertation, machine learning approaches will be explored in various ways to solve problems with DN performance. A new data-driven method is proposed to generate pseudo-measurements from limited available measurement and temporal and meteorological information in an ADN. In order to take advantage of traditional and smart measurements toward grid modernization, a new method based on online learning is applied to combine different measurements at multiple time scales. Moreover, a new method based on a machine learning approach is proposed to perform SE calculations and FDIA detection simultaneously to improve the accuracy and decrease computation time in a DN. The results of the proposed methods show that DN performance in terms of accuracy and effectiveness is improved.

CHAPTER TWO

LITERATURE REVIEW

A power grid consists of three main sections: energy generation, transmission, and distribution networks. The electric power output at the generation side is transmitted over long distances over high-voltage transmission lines [1]. The voltage level is stepped down at the distribution side of the network, allowing the adoption and use of electricity by customers [2]. A typical distribution network (DN) consists of a high-to-medium voltage (HV/MV) transformer and a medium-to-low voltage (MV/LV) transformer [3]. Industrial consumers are connected at the MV grid while residential and commercial customers are typically connected to a LV grid, feeding from the MV/LV transformer [4]. Low-voltage DN density has been getting more crowded over time because of rapidly growing electricity demand [5], [6].

Power networks are also evolving from passive DNs with unidirectional communication paths to ADNs with bi-directional communication. DNs without any DGs are known as passive since the power is supplied from the main grid. In contrast, ADNs are supplied by DGs in addition to a main grid. Moreover, the inclusion of other new features—including penetration of different types of DGs, Electric Vehicles (EVs), variable loads, installation of smart meters and sensors, etc. [7], [8] — convert a traditional DN to smart one.

In Figure 2.1, an overview of an advanced smart power grid structure is shown. As it is clear, the communication among the different sections and the control center are provided by the existence of two-way communications in the smart grid [9], [10]. Smart grids introduce several benefits in different aspects to the existing traditional grid.

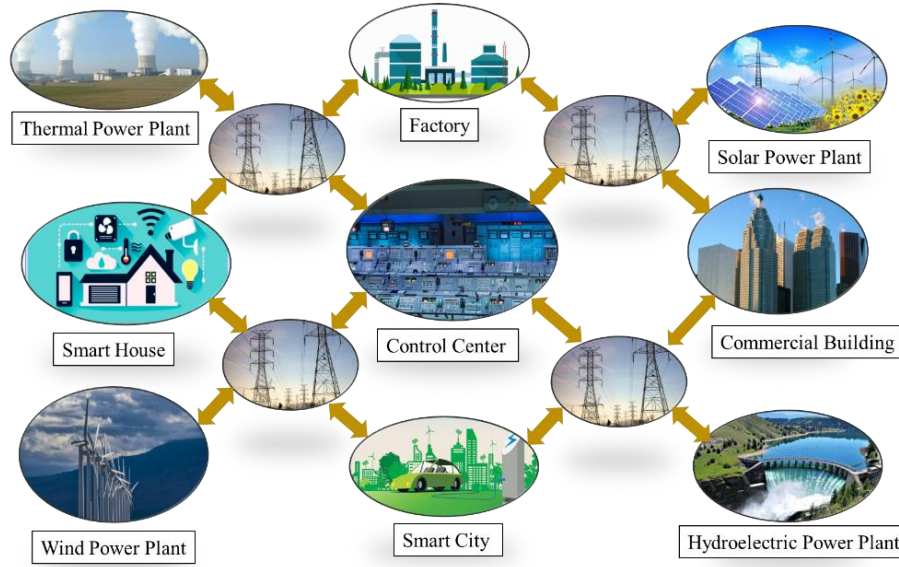


Figure 2- 1. Smart power grid architecture.

These benefits could potentially be promoting renewable and green energy as an alternative source. Developing infrastructure for power storage and electric vehicles could improve the quality and reliability of a power grid. Moreover, these new technologies have resulted in economic and environmental advantages and stimulating the development of smart cities to reduce the energy cost and peak demand.

Even though the integration of these new technologies has paved the path towards grid modernization, the management and control strategy for DNs have become more challenging and complicated [11], [12]. The requirement of observability assessment and need for Distribution System Operators (DSOs) to monitor and control ADNs are intensified by the emergence of these technologies [13]. DSSE algorithms have been introduced to monitor and determine modifications along with the advanced changes in ADNs [14], [15].

SE techniques provide a link between measurement data and control centers to present and monitor all the electrical quantities in ADNs. Transmission System State Estimation (TSSE)

methodologies have also been developed and implemented in TNs from many years ago [16], [17]. Unfortunately, these methods cannot be applied directly in DN [18], [19]. Transmission and distribution networks differ significantly in the following cases: 1) measurement instruments are usually installed at a limited bus, unlike the transmission ones that all buses are monitored by measurements devices. 2) DN are mostly configured in radial or weakly meshed networks, but the topology of TNs is highly meshed. 3) the high number of buses and imbalanced configuration in DN make them different from the balanced structure of transmission grids. 4) due to the low X/R ratios in the DN, decoupled SE algorithms, commonly used for TNs, are not possible in the DN. Therefore, new algorithms should be developed using DN constraints and needs [20]–[22].

The first DSSE method was reported in 1990 when supervisory control and data acquisition (SCADA) systems were introduced [23], [24]. Data from measurement instruments is one of the main requirements to monitor and control ADNs using DSSE calculations. However, measurement device placement is limited and challenging in a DN since it consists of several buses and installation of measurement devices at all buses are prohibitively expensive [25], [26]. The lack of measurements causes issues with observability and the ability to perform DSSE computations.

To address the lack of measurements in ADNs, techniques for generating pseudo-measurements from a limited measurement and historical load and generation data based on Advanced Metering Infrastructure (AMI) measurements have been developed as a crucial task in recent years [27]. However, the quality of the pseudo-measurements is low in both speed and accuracy [28]. Gaussian distribution method is commonly performed to generate pseudo-measurement from load profiles information. It is worthwhile to mention that monthly, weekly,

or daily load profiles could be considered as a part of load profiles information. Since more dynamic behaviors are caused by the new features such as distributed generations, variable loads, electrical vehicles, etc. the conventional methods are not appropriate to generate pseudo-measurements in ADNs. Conventional measurements from Control and Data Acquisition (SCADA) measurements and pseudo-measurements along with physical parameters of a grid are used to perform DSSE calculation in a low sampling rate [29].

SCADA measurements are asynchronous and available with a long interval time and low accuracy [30], but they are much higher than the pseudo-measurement [31], [32]. So, the DSSE results from SCADA and pseudo-measurements are not precise enough to monitor and track changes in an advanced DN.

Installing smart measurement devices, such as Phasor Measurement Units (PMUs) and micro-PMUs (μ PMUs), are currently involved in DN modernization projects to improve the performance of DSSE calculations [33], [34]. PMUs are extensively used in TNs [35], however because DNs have different characteristics than TNs, PMUs are not as commonly installed in DNs [36]. Since PMU measurements are available in synchronous form and can achieve high sampling rates, the incorporation of PMU measurements along with traditional measurements into SE calculations improves DSSE results, and increases the observability, resiliency, and reliability levels of DNs [37]–[39]. However, replacing conventional measurement devices with advanced tools is not possible over a short period of time and it may take several years to modernize a DN.

Because of the presence of both traditional and advanced measurement data during the modernization process, managing, controlling, and performing SE calculations can be

challenging [40]. One of the important issues in the grid modernization process is how different measurements can be incorporated correctly in a control center to perform SE calculations accurately in high-speed operation [41]. Inconsistent reporting rate of measurement devices, and different types of measurement information with various precision levels are two of the most critical problems to perform hybrid SE calculations from traditional and advanced measurement devices [42].

PMUs commonly provide synchrophasor measurements at a rate of 30 samples per second as compared to SCADA measurements which get updated every 2-6 seconds [43], [44]. Hybrid SE methods have been established and developed in power networks to integrate and take the potential benefits of traditional and advanced measurements [45], [46]. Different hybrid TSSE methods to combine traditional and advanced measurements are discussed in [47]. Hybrid methods will be helpful during the power grid modernization development while replacing all-traditional with all-advanced measurement devices [48].

DSSE results not only depend on the availability of real measurements but also the accuracy of real measurements obtained from measurement devices [49], [50]. However, the progress of information and new technologies in distribution grids increase the possibilities for attackers to manipulate information obtained from measurement devices [51], [52]. The accuracy of DSSE results can be diminished significantly when False Data Injection Attacks (FDIAs) affect measurement data [53], [54]. The goal of attackers is to falsify the data and measurement information in power networks [55]. Therefore, DSSE results can be influenced by FDIAs due to dependencies of SE calculations on measurement information [56]. DSSE results are submitted to the control center for further processing including contingency analysis, power flow study,

distribution grid stability and reliability, distribution energy market for energy management, etc. [57], [58]. If FDIAs cannot be successfully detected, the control center performance will be severely degraded, and the likelihood of both economic and physical risks may arise from a wrong decision in a control center [59].

Nowadays, Machine Learning (ML) approaches have been applied widely in DNs to extract valuable data collected from measurement devices and to enhance the DN performance [60], [61]. ML is an evolving branch of computational algorithms and statistical models that computers use to learn and act in an independent manner without human intervention [62]. ML has been developing in recent years for different purposes in solving complicated and challenging problems [63], and the performance and prediction capabilities of ML algorithms change over time as new data is collected and analyzed [64]. ML approaches are classified into four main categories: supervised, unsupervised, semi-supervised, and reinforcement learning [65]. ML techniques are applied widely in DNs for different purposes, such as generating pseudo-measurement, DSSE calculation [66], [67], FDIAs detection [68], [69], DN protection, hybridization of conventional and smart measurements, reliability and resiliency analysis, etc.

Machine Learning (ML) Approach

A significant amount of data is being produced as part of modern power equipment and grid modernization procedures [70]. This data can come from various locations in the power grid, from generation sources to customer loads [71]. Both the data and the advanced power grid facilities can help better realize and analyze power system performance [72]. ML techniques have been widely used and developed in power grids to extract information obtained from different power system sections and components to properly manage and control a power grid

[73], [74]. Forecasting analysis can be performed using model-based and ML-based methods in a power grid [75], [76]. However, various benefits could be obtained using ML approaches rather than using model-based methods [77], [78]. Some benefits are: 1) more accurate results, 2) fast decision-making, 3) the ability to deal with massive data volumes at high sample rates, 4) simultaneous performance of multiple tasks with different levels of uncertainty, and 5) the ability to predict future system conditions and states.

ML methods are advanced mathematical techniques that attempt to mimic human decision-making by using a data-based learning procedure [79]. ML techniques are commonly divided into four categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. These methods are illustrated in Figure 2.2.

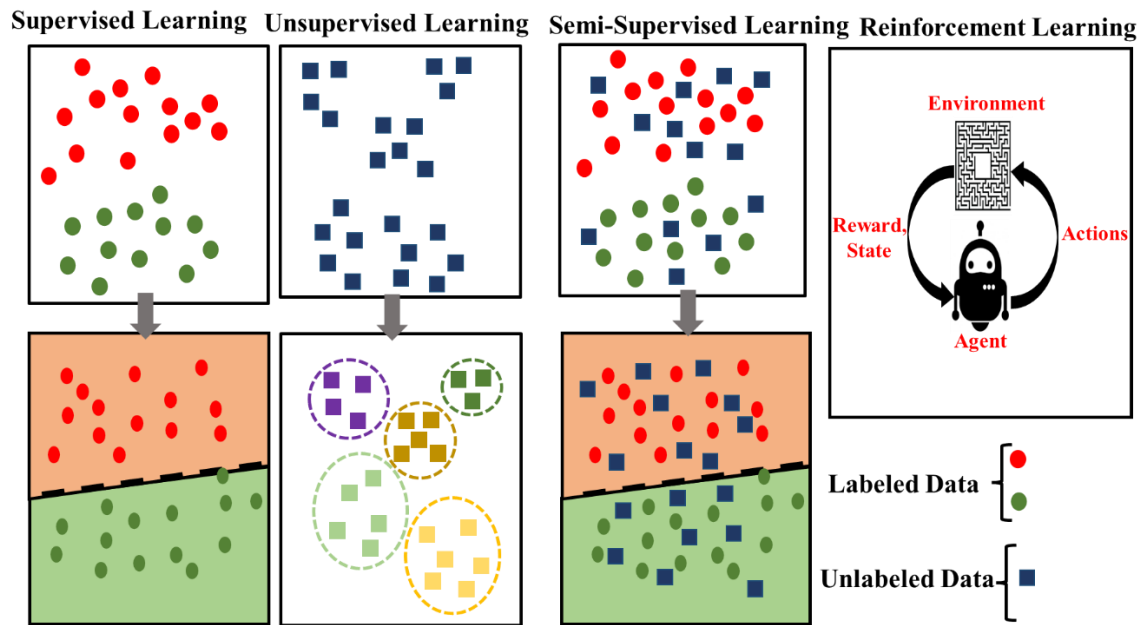


Figure 2- 2. Visual representation of four ML categories: Supervised, Unsupervised, Semi-Supervised, and Reinforcement.

In supervised learning, a corresponding output (Y) is available for each input data (X). The relationship between X and Y is represented as $Y = F(X)$. Supervised learning algorithms estimate the function F , which can be used to predict class labels or make decisions for unseen examples [80]. In unsupervised learning, the input dataset is not labeled (i.e., the corresponding output is not available), and a function tries to find the hidden patterns in the training dataset [81]. In semi-supervised learning, some of the given data is unlabeled, but the training process can take advantage of the points that do have labels [82]. Reinforcement learning is a computational approach where the algorithm (agent) learns to achieve a goal in an uncertain, potentially complex environment [83]. Reinforcement learning does not need labeled data for a training procedure, but there needs to be a way for the agent to receive positive or negative feedback after making a decision [84].

Supervised Machine Learning Method

Regression and classification are two common subcategories of supervised machine learning (SML) [85]. Regression and classification methods can be carried out using various techniques, some of which are shown in Figure 2.3.

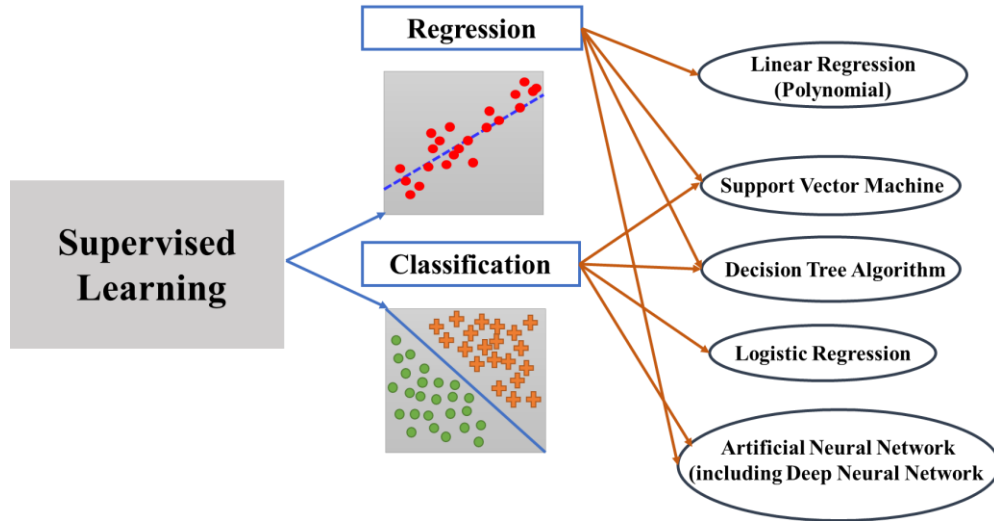


Figure 2- 3. Different SML techniques, categorized by regression and classification.

Regression. A regression problem is applied to predict continuous output values (i.e, targets) based on the available input data (i.e, independent values) [86]. The main objective of a regression method is to discover a relationship between input variables (features) and target values. This algorithm is trained based on historical data to predict or estimate unseen future data. There are several types of regression algorithms, including linear regression [87], polynomial regression [88], support vector regression [89], decision tree regression [89], and artificial neural network (including deep neural network) regression [90].

A regression problem is applied to predict continuous output values (i.e, targets) based on the available input data (i.e, independent values) [86]. The main objective of a regression method is to discover a relationship between input variables (features) and target values. This algorithm is trained based on historical data to predict or estimate unseen future data. There are several types of regression algorithms, including linear regression [87], polynomial regression [88], support vector regression [89], decision tree regression [89], and artificial neural network (including deep neural network) regression [90].

Classification A classification problem is applied to predict the category or class label of a given input based on its features [90]. The output of a classification model is a discrete class label or a probability distribution overall class label. Classification problems with two classes are known as binary classification problems; problems with more than two classes are known as multi-class classification problems [91]. In ML classification, a model is learned using the values and labels of a set of training data, which is used to predict unseen future data [92]. There are several types of classification algorithms, including logistic regression [93], support vector classifier [94], decision tree classifier [95], artificial neural network (including deep neural network) classification [96].

A classification problem is applied to predict the category or class label of a given input based on its features [90]. The output of a classification model is a discrete class label or a probability distribution overall class label. Classification problems with two classes are known as binary classification problems; problems with more than two classes are known as multi-class classification problems [91]. In ML classification, a model is learned using the values and labels of a set of training data, which is used to predict unseen future data [92]. There are several types of classification algorithms, including logistic regression [93], support vector classifier [94], decision tree classifier [95], artificial neural network (including deep neural network) classification [96].

a) Linear Regression (LR)

LR is applied to predict the dependent variable (y) based on one or more independent variables (x) [97]. The simple form of linear regression can be represented as the following equation:

$$y = w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n, \quad (1)$$

where y is the target value, w_0 is the y -intercept, $w_1, w_2, \dots, w_n$ are weights associated with each input variables, and $x_1, x_2, \dots, x_n$ are the input variables (features). In LR, the main purpose is to estimate the coefficients $w_0, w_1, \dots, w_n$ that minimize the error (e) between the predicted values ($\hat{y}$) and actual values (y). This is typically done using the least square approach where the minimization problem is formed as:

$$\sum e_i^2 = \sum (y_i - \hat{y}_i)^2. \quad (2)$$

b) Polynomial Regression

Polynomial regression is an extension of linear regression where the relationship between the independent variable (x) and dependent variable (y) is represented by a polynomial function of x [98]:

$$y = w_0 + w_{11}x_1 + w_{12}x_1^2 + w_{13}x_1^3 + \cdots + w_{1n}x_1^n + w_{21}x_2 + w_{22}x_2^2 + w_{23}x_2^3 + \cdots + w_{2n}x_2^n + \cdots + w_{mn}x_m^n, \quad (3)$$

where y and x are dependent and independent variables, respectively. $w_0, w_{11}, \dots, w_{mn}$ are coefficients that determine the shape of the polynomial curve. The degree of the polynomial is n , and this represents the maximum complexity of the curve. Polynomial regression can be useful when the relationship between the variables shows a curved pattern, and it allows to capture the non-linear relationship between the features [99].

A simple example of linear and polynomial regression models is shown in Figure 2.4.

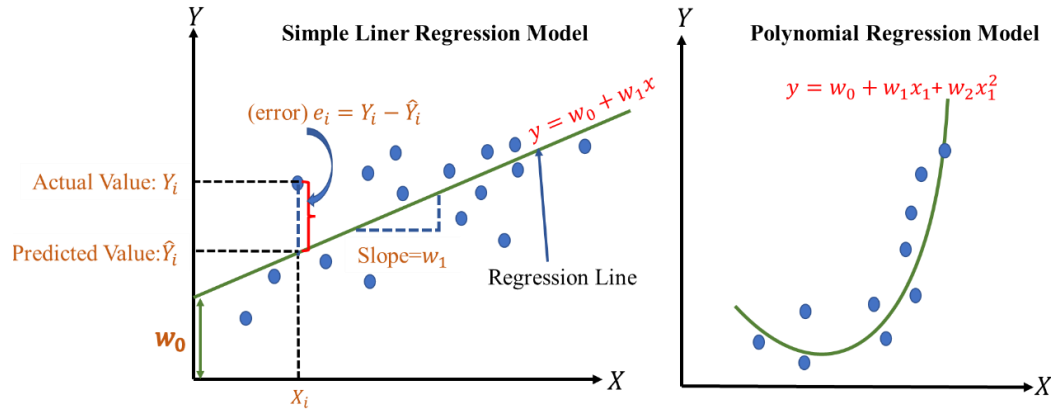


Figure 2- 4. Linear and polynomial (2nd order) regression models for 1-dimensional inputs.

c) Logistic Regression

Logistic regression is a common classification approach that is applied to predict both binary and multi-class outcomes [100]. The relationship between a dependent variable (target) and one or more independent variable (feature) is modeled using class probability estimation. As it shows in Figure 2.5, logistic regression model uses a logistic or sigmoid function to transfer the features (independent variables) into a value between 0 and 1.

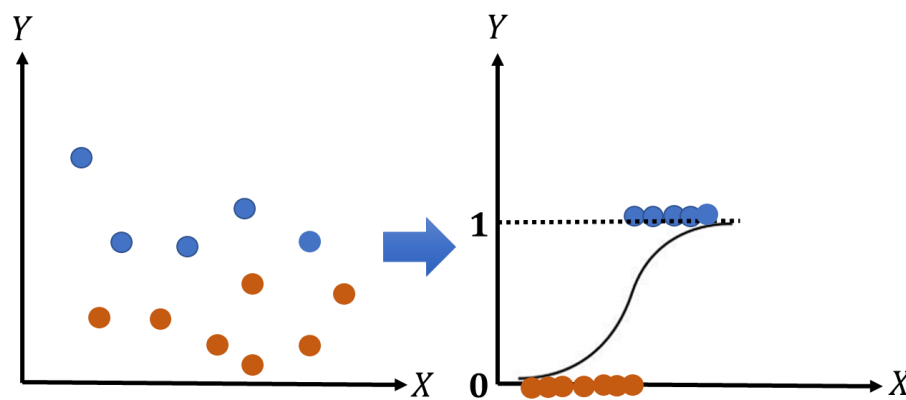


Figure 2- 5. Logistic Regression Model.

Simplicity, the interpretation of model coefficients, and accurate and efficient training are the advantages of logistic regression algorithm [101].

d) Support Vector Machine

A Support Vector Machine (SVM) is a SML technique which can be applied for classification and regression tasks [102]. Support Vector Regression (SVR) is utilized for regression tasks where the target variable is continuous, while Support Vector Classification (SVC) is applied for classification tasks where the target is categorical. The main objective in SVR is to find a hyperplane that best fits the training data while maximizing the distance (margin) to the closest training data [103], as shown in Figure 2.6(a). The hyperplane reflects the regression function that can be used to predict unforeseen future data. Those data points that lie closest to a given hyperplane are applied to determine the position of the hyperplane. The tolerance region around the regression line is defined as ε and allows some degree of flexibility in the placement of the learned hyperplane. The SVR can be kernelized to estimate a nonlinear regression curve [104].

SVC is a popular approach which can be adapted mostly in a binary classification task [105]; it can be implemented in both linear (and non-linear, using kernels) classification by finding an optimal hyperplane that maximize the margin between the class boundary [106]. As shown in Figure 2.6(b), data points which are close to the hyperplane are called support vectors and the position and orientation of the hyperplane are determined by support vectors. SVC might not be suitable for large-scale dataset due to the computational complexity [107].

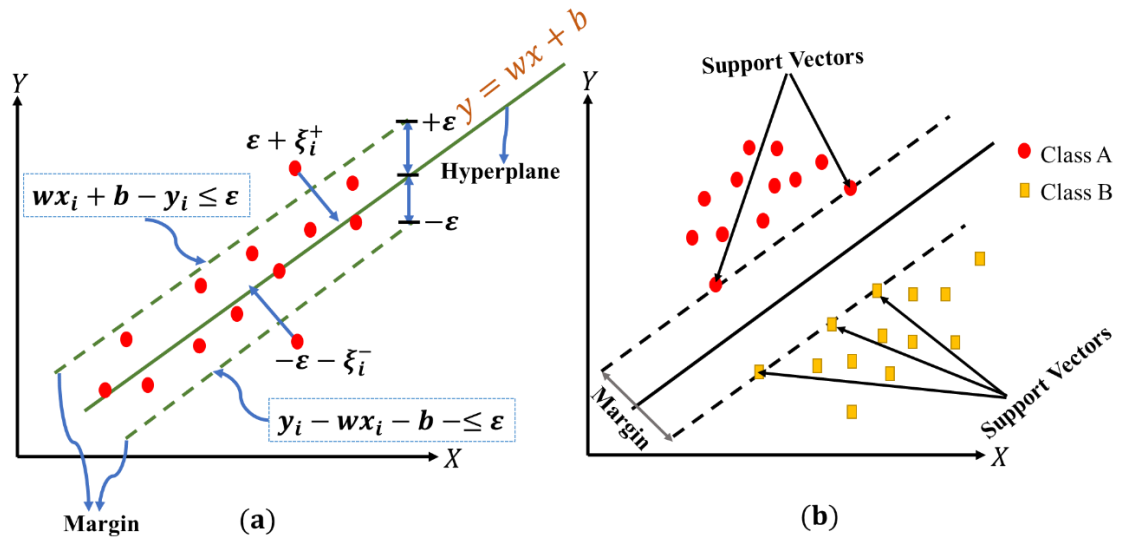


Figure 2- 6. (a) Support vector Regressor; (b) Support Vector Classification.

e) Decision Tree

The Decision Tree is a popular and intuitive ML model applied for classification and regression analysis [108], [109]. Figure 2.7 illustrates a simple decision tree model that consists of an internal node and leaf. A decision is created based on a specific feature, and an output is predicted at each leaf node of the tree. A Decision Tree Regression (DTR) model has a broad range of applications, especially when the relationships between independent and dependent variables are nonlinear or when model interpretability is important [110]. However, the nature of the decision tree results in highly quantized regression values as the output. A Decision Tree Classification (DTC) algorithm can be applied to both binary and multi-class classification problems [111]. However, Decision Trees are susceptible to overfitting if the tree is allowed to grow deeply or the data is noisy [112].

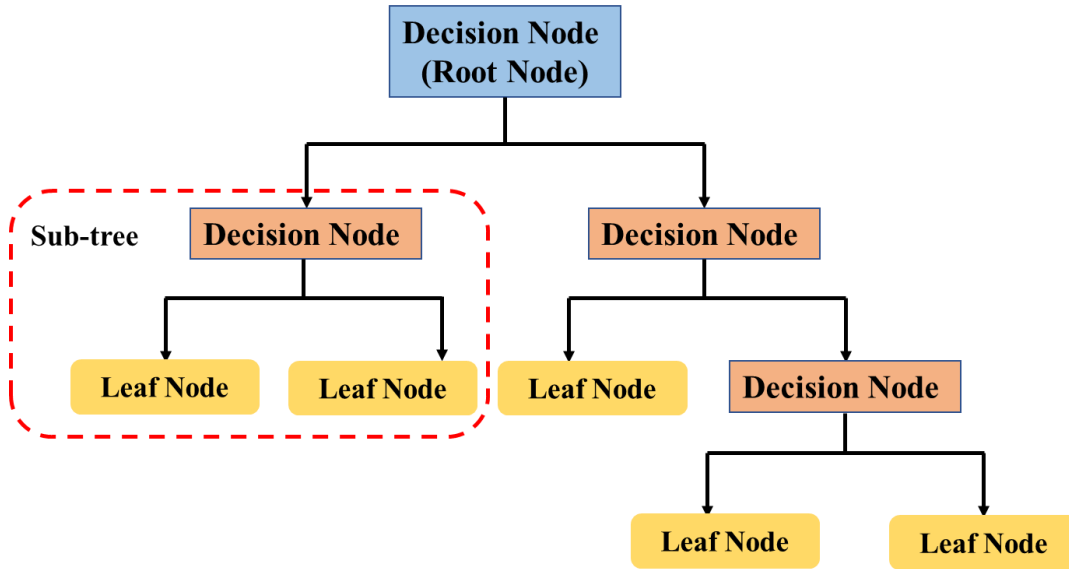


Figure 2- 7. Decision tree algorithm

f) Deep Neural Network

As shown in Figure 2.8(a), a single-layer perceptron includes one input layer with one or many input units and an output layer with one or many output units [113]. In contrast, a multilayer perceptron has a number of hidden layers as well. DNN is a part of ML and special type of Artificial Neural Network (ANN) inspired by the human brain. A DNN is generally considered as having 2 or more hidden layers [114], [115]. In feed-forward DNN models, the information flows in only one single direction: from the inputs, through hidden layers, to the output layer.

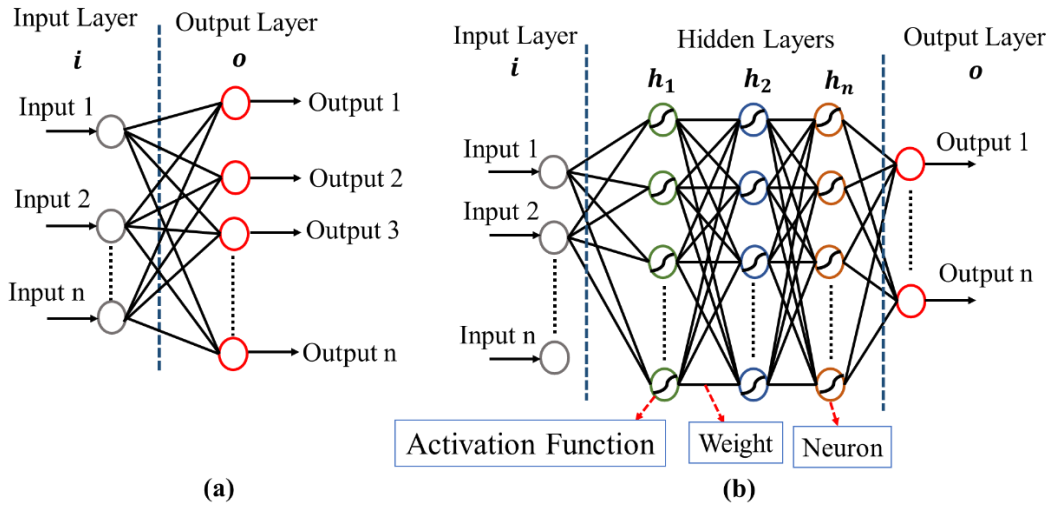


Figure 2- 8. (a) A single-layer perceptron; (b) A multi-layer perceptron (DNN)

Figure 2.8(b) shows a DNN model consisting of an input layer, activation functions, hidden layers, weights, neurons, and an output layer. The input data is received by neurons in the input layer and passed (transferred) to the different layers in the network. The hidden layers are located between the input and output layers; the number of hidden layers and the number of neurons are determined by means of empirical testing or trial and error [116]. In order to predict an accurate output based on the original input, the weights of the DNN are constantly updated during training. In classification or regression problems, a DNN model is applied to predict a class label or a numeric value, respectively [117]. The weight on the connection shows the magnitude between two connections and falls within the range of 0 to 1. The sum of the weighted inputs is determined by the neurons; moreover, another functionality of neurons is to initiate an activation process to normalize the sum of the weighted inputs. The activation function is applied within the hidden and output layers and can be either linear or nonlinear. Sigmoid, tanh, and Rectified Linear Unit (ReLU) are three common forms.

Assessment of Regression and Classification Method

Different criteria are applied to evaluate the performance of regression and classification SML methods [118]. These criteria are considered to determine the most effective method for a specific task.

a) Evaluation Metrics in Classification Method

Several metrics can be used to evaluate the classification technique performance [119]. Some common evaluation metrics for classification methods are reviewed:

- Accuracy:

This is the most common evaluation metric in classification problems. Accuracy is defined as the ratio of the number of correct predictions over the total number of predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

TP stands for True Positive, where both predicted and actual outputs are positive. TN stands for True Negative, where both predicted and actual outputs are negative. FP stands for False Positive, where predicted and actual outputs are positive and negative, respectively. FN stands for False Negative where predicted and actual outputs are negative and positive, respectively.

Note: The following evaluation metrics are commonly used for a binary classification problem, but it could also be developed for multiclass classification model [120].

- Precision:

This is defined as the proportion of correctly predicted positive output to the sum of number of true positive plus false positive.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

- Recall:

This is defined as the proportion of correctly predicted positive output to the sum of the number of true positive plus false negative.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

- F_1 Score:

Precision and recall are combined into a single metric in the F_1 Score, which conveys the balance between the precision and recall. F_1 Score is particularly useful when there is an imbalance between classes in the dataset.

$$F_1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (7)$$

b) Evaluation Metrics in Regression

Different evaluation metrics can be implemented to analyze regression model performance [121]. Some common evaluation metrics for regression methods are as follows:

- Mean Square Error (MSE):

This is a popular error metric for the regression problems. The MSE is calculated as the mean or average of the squared differences between predicted and estimated output values of a dataset; it is defined as:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (8)$$

- Root Mean Square Error (RMSE):

It is an extension form of the MSE. RMSE can be calculated as follows; its unit is the same as the original unit of the target value.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} = \sqrt{MSE} \quad (9)$$

- Mean Absolute Error (MAE):

It is calculated as the average of the absolute error values:

$$MAE = \sqrt{\frac{1}{N} \sum_{i=1}^N abs(y_i - \hat{y}_i)}. \quad (10)$$

where y and $\hat{y}$ are the estimated and predicted values in the dataset, respectively, and N is the sample size. Like RMSE, the unit of the MAE and the predicted value are the same.

Machine Learning Approach for Generating Pseudo-Measurements

DSSE calculations are strongly dependent on real measurements collected by various instruments installed on the DNs [122], [123]. Measurement devices are commonly installed in the HV/MV primary DNs. But the placement of measurement devices on the LV side is limited and challenging since LV substations consist of several LV lines and installation of measurement devices at all buses are prohibitively expensive in the near future [25], [26]. To address the lack of measurements and improve the quality of DSSE calculations in DNs, generating pseudo-measurements (corresponding to off-line load estimation) from historical data and load forecasting techniques are established in DNs [28], [124]. Monthly, weekly, or daily load profiles along with other data can also be utilized to generate pseudo-measurements. Gaussian distribution method is commonly performed to generate pseudo-measurement from load profiles information. Gaussian Mixture Model (GMM) is proposed in [27] to generate pseudo-measurements in ADNs. Other conventional methods, including stochastic time series and regression-based methods, are also used to predict energy consumption [125]. This classical method performs well in small-sized power networks, but it is not very effective in active large-sized networks.

Moreover, solar and wind energy systems, which are commonly used in ADNs, are variable energy sources and add more uncertainties to ADNs. This limits the application of conventional methods for generating pseudo-measurements in ADNs [126], [127]. So, in recent years, ML approaches have been introduced to obtain pseudo-measurements from available data to address some problems with the conventional methods [128]–[131]. Methodologies for generating pseudo-measurements methods are widely utilized by Distribution System Operators (DSOs) to perform SE calculation in order to improve the performance of DN effectively. Recent advancement of applying data-driven approaches in DNs makes a significant improvement to generate pseudo-measurements from a limited available measurement.

In [132], a gradient boosting tree model is proposed to generate pseudo-measurements in ADNs. Authors in [133] apply Relevance Vector Machines (RVMs) [134] to perform pseudo-measurement modeling in distribution networks. An Adaptive Nonlinear Auto-Regressive exogenous (NARX) is used to perform load forecasting in DNs in [135]. ANN and DNN are also widely used to generate pseudo-measurements in DNs. In [131], an ANN is applied to generate active and reactive power injections from a limited real measurement. In [136], a distributed game theoretic based on SML is applied to generate active power injection using AMI data. In [137], multi-timescale learning structure based on ANN model is developed to generate an active power injection for commercial, residential, and industrial sections from monthly billing data. In [138], an ANN, linear regression, and Support Vector Regression (SVR) are applied to generate load profile and electricity charges of consumers. In [139], an ANN based pseudo-measurement generating algorithm (PMG-ANN) is applied to address a lack of real measurement and low-observability issue in the DN.

In [140], a DNN approach is proposed to identify consumption behavior and load forecasting. A novel pooling-based DNN model is proposed in [141] for household load forecasting.

Machine Learning Approach for Hybridization Traditional and Smart Measurements

Combining multiple measurements at different reporting rates to perform precise SE calculations in a control center is identified as one of the notable topics in the grid modernization efforts. To execute hybrid SE calculations based on smart and conventional measurements, both advanced measurement data and traditional measurement data must be incorporated. Advanced measurement data has high precision and high resolution and is collected at a rate of up to 30 samples per second. On the other hand, traditional measurement data has lower precision and resolution and is updated every 2-6 seconds [123].

Machine learning methods have also been developed to hybrid traditional and smart measurement data for DN modernization. In [142], different attributes of PMU and SCADA measurements are determined based on data accuracy, time scale, and refreshing frequency. Then, a specific solution is created based on the differences between measurement properties.

In [143], in order to include both SCADA and PMU measurements for DSSE calculation, the state estimator switches between a Weighted Least Square (WLS) and Weighted Least Absolute Value (WLAV) calculations.

In [144], a new hybrid SE method based on Particle Swarm Optimization (PSO) [145] and Artificial Neural Network (ANN) approaches is proposed to combine the different measurements.

In [146], nonlinear static state estimation, linear static state estimation, and linear dynamic state estimation are applied to hybrid SCADA, PMU and AMI measurements in ADNs.

In [147], a new hybrid state estimator considering distributed state estimation method is proposed to use SCADA and μ PMU measurements in an unbalanced DN.

In [148], the improved robust and linear state estimators are applied to respond to different frequency levels on measurement data from SCADA and PMU for medium voltage DNs.

In [149], a hybrid simultaneous state estimator based on WLS method is proposed to incorporate both SCADA and PMU measurements.

In [150], a hybrid SE method based on WLAV estimator and Deep Neural Network (DNN) methods is proposed to manage the unsynchronized AMI and SCADA measurements and monitor an ADNs.

In [151], SCADA, μ PMU, and Smart Meter (SM) data are used for an unbalanced DNs, where Kalman smoother and Expectation Maximization (EM) methods are used to manage the missing measurements and combine measurements in different time samples, respectively.

In [152], a static state estimator is used to perform DSSE calculation using both SCADA and PMU measurements. When SCADA measurements are not available, a fast state reconstruction algorithm is applied using only PMU and pseudo-measurements.

In [153], a spatio-temporal estimation generative adversarial network (ST-EGAN) is proposed to generate pseudo-measurements to perform DSSE calculations in high resolution when SCADA and PMU measurements are not available simultaneously.

In [154], an improved sequential state estimation method is proposed to take advantage of AMI and (μ PMUs) measurements to address the asynchronization issue for performing DSEE calculations.

Machine Learning Approach for Distribution System State Estimation

State estimation (SE) calculation has been introduced to monitor the performance and secure operation of power grids [155]. The SE process is used to estimate the power system states based on available measurements and network parameters. SE computation is considered the core of Energy Management System (EMS) since SE results are applied to control and improve the operation performance of a power grid [156]. Contingency analysis, security assessment, economic dispatch, and other functions are possible using SE results [157]. TSSE has been successfully developed and applied to TNs for several years. Transmission grids are mainly designed in mesh configuration and found as an over-determined and fully observable system since the number of available measurements are more than the number of state variables in a TSSE calculation [158]. However, DNs are configured with either a radial or weakly meshed structure and have a fairly large number of nodes. This makes it impossible to place measurement devices at all nodes and leads to the lack of measurements for DSSE calculations [159]. Therefore, because of the main differences between transmission and distribution networks, TSSE methods are not directly applicable in DNs and new methods and approaches should be designed to address the requirements of DNs[160].

DNs are generally responsible for supplying high-quality end-customer service and can be used for operational and management planning. These are critical tasks, which can be accomplished using appropriate DSSE calculations [161]. DSSE computations are highly

dependent on measurements which are obtained from measurement tools. However, real measurements are not sufficient to perform an accurate DSSE calculation [162]. Therefore, generating pseudo-measurements from an available measurement and historical data is widely applied to address the lack of measurements and execute DSSE computations precisely [163]. The state vector of the power grid can be defined as a set of variables; when the state variables are calculated, other electric power quantities can be computed from these states [164],[165]. The state vector for a DN can be defined as the voltage magnitude and phase angle at each bus or the branch current magnitude and angle of each branch in a DN [166].

Model-based DSSE techniques based on measurement information and network parameters have been established to acquire state vectors. Some traditional model-based DSSE techniques and their advantages and disadvantages include [167]:

- Weighted Least Square (WLS) is applicable in both transmission and distribution networks; it is a fast and straightforward method, but it is sensitive to bad data.
- Least Median of Squares (LMS) is robust against bad data, however it requires high computational complexity and measurement redundancy.
- Least Trimmed Squares (LTS) is robust against bad data, but both computational complexity and memory utilization are high.
- Least Absolute Value (LAV) is robust against bad data and line impedance uncertainty, but in contrast, it is sensitive to measurement uncertainty and leverage point.

DSSE calculation has received much attention due to the new features of an advanced DN including DGs, Electric Vehicles (EVs), two-way communications, smart buildings and industries, energy storage systems, and smart meters [168]. A DN has become increasingly

complex by these advanced features and new DSSE techniques are required to improve SE results in terms of accuracy and speed [169]. Thus, DSSE approaches based on ML techniques are developed widely to enhance SE results since these algorithms are mainly established based on available data, unlike model-based methods which rely on system information as well [170].

With the development of sensor and smart meter technologies, the direction for further research on DSSE techniques using ML methods in DNs has become important among the researchers [171]. In [172], two DNNs are used to detect topology and perform DSSE calculation. Switch status is identified by a DNN model using only sparse measurement data. Another DNN is trained to perform DSSE calculation for an unbalanced three-phase DN. Voltage magnitudes and phase angles are defined as the state vector. The authors tested and verified their proposed method for both radial and meshed networks with Photovoltaic panels (PVs). Moreover, non-Gaussian noise is considered by the authors for a smart meter measurement to show the robustness of the method.

In [173], an ANN-based model is used to perform DSSE calculations using a limited number of measurements in the presence of two types of DGs (PVs and wind turbines). Voltage magnitude and branch currents are defined as state variables. The method is tested and verified on a benchmark DN and a real DN considering high penetration of DGs. They compared the accuracy of the method with the WLS approach using the same measurement data (and pseudo-measurements) and showed that the method works well with high accuracy and fast execution time.

In [174], a new DNN-based method named pruned physics-aware neural network (P2N2) is used to improve voltage magnitude estimation. The connections between hidden layers are

determined based on the physical topology of the DN. They utilized one year load data to generate different scenarios using power flow calculations. These data along with PMU measurements are utilized to train and test a method under different observability conditions. They compared the method with the WLS approach and reported that the P2N2 method has better performance in terms of estimation accuracy in a low observability condition.

In [175], the computational complexity of linear and nonlinear WLS methods and ANN models are compared in Floating Point Operations (FLOPs). This technique is a machine independent measure to calculate the run time of each method [176]. It is shown that the ANN method has a faster running time than the WLS methods. The paper compared FLOP values with the elapsed time [177], which applied in many studies, and showed that elapsed time may not work well in a small DN, whereas the FLOP criteria is applicable in different DN sizes.

In [178], a data driven approach based on an ANN model is used to perform distributed state estimation using limited PMU measurements. Voltage magnitudes and phase angles are defined as the state vector. A DN is divided into several areas, and an ANN method is applied to obtain SE values for each area for each phase ($\emptyset = \{A, B, C\}$), then SE results from each area are sent to a control center for further processing and a global DSSE calculation is performed based on ANN model to obtain the final results. Multi-area data-driven and model-based approaches based on the WLS method are compared in terms of accuracy and execution time, and it is shown that state estimation variables from a data-driven approach follow a real value from a power flow calculation in a high accuracy at a high speed compared to WLS method. Distributed state estimation methods are introduced in a DN to address a lack of measurement and improve SE results [179]. In distributed state estimation approaches, a DN is divided into several areas

based on different criteria and a state estimation calculation is performed in each area separately. Then, state estimation results from each area send to a control center for further processing and obtain final state estimation results [180].

In [181], a DN is divided into several independent sub-areas, and a certain ANN is designed for each area based on an available measurement in each area. A method is tested and verified at one LV and one MV distribution networks.

As mentioned, unlike model-based methods, data-driven approaches do not depend on system information. However, data-driven methods require a certain amount of data to effectively perform DSSE calculations. So, some studies propose a hybrid method based on data-driven and model-based approaches in order to take advantage of the strengths of each method [182].

In [150], in order to utilize different measurement data at different time scales, a hybrid method based on DNN and Weighted Least Absolute Value (WLAV) is implemented. When only SCADA (Supervisory Control And Data Acquisition) measurements are available, a DNN model is used to perform DSSE calculation; when SCADA and AMI (Advanced Metering Infrastructure) measurements are available, WLAV method is applied to obtain SE results.

In [183], distribution learning and regression methods are applied to generate power injection in a DN and a DNN model based on the Bayesian DSSE approach is used to obtain SE results. Voltage magnitudes and phase angles are defined as the state vector. Their method is tested and verified in both a radial and meshed network, showing that their model outperforms other DNN models.

In [184], a shallow ANN model is used to calculate an initial value using historical load and generation data, pseudo-measurements, PMU measurements, and current magnitude measurements. The initial values are used to perform SE calculations based on Gauss-Newton iterative estimation [185]. The method shows acceptable performance in terms of accuracy and execution time.

In [186], historical data of DPMU (Distribution Phasor Measurement Unit) measurements (voltage and current magnitude and phase angle) and limited smart measurement data are utilized to train a feed-forward neural network to estimate the voltage magnitudes and phase angles of a DN.

In [187], a novel ANN model is developed to calculate state variables with high accuracy under different scenarios including topology changes, load variation, and bad data in a DN. Their method is also applicable to predict measurements such as power flows and power injections. The proposed method is tested in different distribution grids with and without DGs to show its robustness.

Machine Learning Approach for False Data Injection Attacks in a Distribution Grid

Distribution System Operators (DMOs) strongly depend on an advanced Distribution Management System (DMS) to further develop smart DNs [188]. DSSE results are known as a fundamental tool for DMS, since most of the DMS functions—including topology analysis, volt/VAR control, distribution load forecasting analysis, distribution congestion management, etc. are completed using SE results [189]. Moreover, additional advanced features are added to a DN as part of grid modernization solutions [190]. Therefore, an accurate DSSE calculation plays a crucial role for advanced DN operation and performance. The accuracy of SE results is directly

affected by measurement data; so with an increase in the installation of smart measurement devices such as PMUs, μ PMUs (micro-PMUs), and AMI (Advanced Metering Infrastructure), DSSE results could be enhanced significantly [191]. However, the possibility of false data injection attacks (FDIAs) on measurement data is intensified by an ongoing increase of measurement devices in a DN; such attacks can degrade the precision of DSSE results [192]. FDIAs are a kind of cyber-attack that is introduced in a DN to falsify measurement data and disrupt a DN's normal operation [193].

The concept of FDIAs in power grids was first studied in 2009 [194]. FDIA detection has been established and developed in DNs to maintain the integrity and accuracy of DSSE results [195]. The purpose of bad data detection (BDD) is to identify the faulty data in measurement information [196]. In a conventional BDD method, the ℓ_2 -norm of the residual measurements is computed based on estimated and real measurements after DSSE calculations to determine if a measurement contains bad data. If the residual value is greater than a pre-defined threshold, bad data is unlikely to exist [197].

In [198], a hybrid Weighted Least Absolute Value (WLAV) method is proposed to use Supervisory Control And Data Acquisition (SCADA) and micro-PMU measurements for three-phase unbalanced DNs. The robustness of WLAV and WLS estimators are compared against potential FDIAs, and it is shown that WLAV has a better performance to enhance the security of the DNs.

In [199], two distributed sparse state estimation and attack detection methods are studied to make a DN observable and perform FDIA detection locally in DNs.

In [200], the interval affine state estimation method is applied to identify attacks in measurement data by considering the upper and lower boundaries of the state variables.

However, since FDIAs can successfully bypass the traditional BDD method, FDIAs cannot be commonly identified by conventional BDD methods and errors could be introduced into a DSEE result.

Therefore, ML techniques are implemented to detect FDIAs efficiently along with applying ML methods to monitor and control DNs [68]. Various types of ML techniques have been analyzed for FDIAs in a DN. FDIA detection is considered a supervised binary classification problem.

In [201], the abilities of different machine learning approaches are tested to identify attacks in ADNs. Furthermore, various scenarios are considered to verify that FDIAs can be identified using machine learning methods.

In [202], a new method based on a SML approach is used to detect FDIAs on control system of a DN in presence of DGs. First, historical voltage and branch current magnitude along with meteorological data (solar irradiance and temperature) are used as inputs in a RF (Random Forest) regressor to estimate voltage magnitude variables. Then, a logistic regression method is applied to detect and locate a FDIA on a control system. A method is tested with both centralized and distributed DN. A 240- bus DN (real-based in the USA), and IEEE 123-bus are used case studies. They showed that FDIAs could be identify with up to 99% accuracy.

In [203], a new method based on a DNN method is used to identify the type and location of FDIAs in an ADN. The method is designed based on a multi-output DNN classification methodology and named as fault and attack location and classification (FALCON). They

compared and analyzed their method with other methods, and it is shown that the FALCON method could detect the type and location of FDIAs on IEEE 13-bus distribution grid with up to 98% accuracy.

In [204], a new DNN-based method is applied to detect FDIAs in a DN. A network training function is updated based on Levenberg-Marquardt (LM) backpropagation [205]. Active and reactive power injection and voltage magnitude and phase angle are considered as measurements which are manipulated by FDIAs. IEEE 33-bus DN is used to validate a proposed method. FDIAs could be identified with high accuracy using this method.

In [206], a SVM approach is applied to identify FDIAs on μ PMU measurements. Different types of attacks are defined and tested on several kernels to find a robust one which is against FDIAs. The authors showed that their method can detect the attack and the number of attacks on a data set.

In [207], errors at different levels, distributions, and biases are considered on PMU measurements to classify events using different SML methods including k-Nearest Neighbor (k-NN) [208], Naive Bayes (NB) [209], Decision Tree (DT), Ensemble (EN) [210], SVM, and ANN. IEEE-123 bus DN is considered to test and verify the presented methods. Their work shows that ANN has the best performance because of its flexibility and robustness against noise and non-linearity behaviors.

Conclusion

In this thesis, different SML techniques are developed to enhance performance and management of ADNs. Pseudo-measurements from a limited real measurement, temporal and meteorological information using a DNN approach are generated in an ADNs considering the

dynamic characteristics of DGs. Then, pseudo-measurements are utilized to perform DSSE calculation to validate the accuracy of pseudo-measurements generated.

A new method based on a DNN approach is applied to perform DSSE calculation and detect FDIAs simultaneously in ADNs considering two types of FDIA. A regression and classification tasks are completed at output layer of a DNN at the same time. The results indicates that a proposed method has a better performance in terms of accuracy and execution time in comparison with a conventional DSSE method.

A new method based on an online machine learning approach is proposed to hybridize traditional and advanced measurements by using advantage of all available measurements to improve DSSE results.

CHAPTER THREE

INCORPORATING TEMPORAL AND METEOROLOGICAL DATA FOR GENERATING
PSEUDO-MEASUREMENTS IN ACTIVE DISTRIBUTION POWER NETWORKS

The measurement and monitoring of power systems are well developed and prevalent for transmission networks, but installing measurement devices is limited on the side of HV/MV distribution networks and commonly absent in LV distribution networks. The accuracy of DSSE results highly depends on the real measurement which are obtained from measurement devices, but the DSSE calculations face a major challenge, as DNs lack real-measurements and have inherent observability issues. Due to economic constraints, installing measurement devices at all buses (to fully monitor a DN) is not feasible in the near future. On the other hand, maintaining the observability and controllability of DNs is becoming more important while advance features are continuously added to DNs. Also, real measurements are not always available due to topology modification, sensor failures, communication errors, cyberattacks, and electrical contingency.

Generating pseudo-measurements from historical data and temporal and meteorological information using either probabilistic and statistical methods or machine learning approaches could provide the solution to address the lack of measurement and maintain the observability and controllability of ADNs. In this dissertation, a new method based on SML is proposed to generate pseudo-measurements in ADNs with a penetration of DGs and a limited number of real measurements.

In recent years, ML approaches have been developed to calculate pseudo-measurements from available data to address the deficiencies of traditional methods [128]. Some of traditional and data-based methods are reviewed in section 2.4.

Deep Neural Networks (DNN) have been applied as a prominent data-based technique to handle the limitation of a dataset by tuning higher hidden layers to generate pseudo-measurements [6],[211]. The performance of a DNN depends on tuning different hyperparameters: number of neurons, number of hidden layers, learning rate, dropout rate, activation function, and more [212]. However, increasing the number of layers and neurons in a DNN cannot always bring an acceptable performance since the network becomes more complex and it needs more data to train. Due to the lack of available measurements for pseudo-measurement generation in ADNs, a DNN approach should be used along with other methods to improve the accuracy of pseudo-measurement results and outperform a DNN performance.

While most previous works improve the accuracy of pseudo-measurement results by implementing ML approaches, they have the following limitations:

- 1- High penetration and dynamic behavior of DGs in clear and cloudy sky conditions are not well studied.
- 2- The lack of measurements is not commonly considered as one of the limitations in ADNs for training an ML model.
- 3- Temporal and meteorological features (seasonality patterns, time-of-day, temperature, cloudiness, etc.) have not yet been considered when addressing the dynamic behavior of a training model for generating pseudo-measurements.

Given the mentioned shortcomings, the main contributions of this work are:

- 1- Different criteria, including temporal and meteorological features, are considered for dividing the training data. Thus, multiple models (base learners) are trained, each tailored to different conditions.
- 2- Designing and developing a system of base learners along with a meta learner generates accurate pseudo-measurements with limited available measurements.
- 3- A single model may fail to obtain an accurate pseudo-measurement result since the large amount of training data is required for effective training. So, it is focused on how an additional method, along with a DNN approach can be applied to address this deficiency and make a training model robust against different dynamic behaviors of a DN.
- 4- The proposed method is compared using different criteria in the case of the training data being divided based on a random split method. We show that by considering temporal and meteorological features, a training model has a better performance.

Generated pseudo-measurements along with available measurements are used to perform DSSE calculations based on the WLS method [160] to assess the accuracy of the proposed method. The effectiveness of the proposed method is evaluated using a modified IEEE standard 69 bus DN considering limited available measurements in presence of DGs.

Methodology

As shown in Figure 3.1, inputs (available measurements, month, day of week, hour of day, temperature data, wind speed, and solar irradiance) are fed into a training model to generate pseudo-measurements. Pseudo-measurements, along with available measurements, are used to perform WLS calculations to obtain the state variables of a DN.

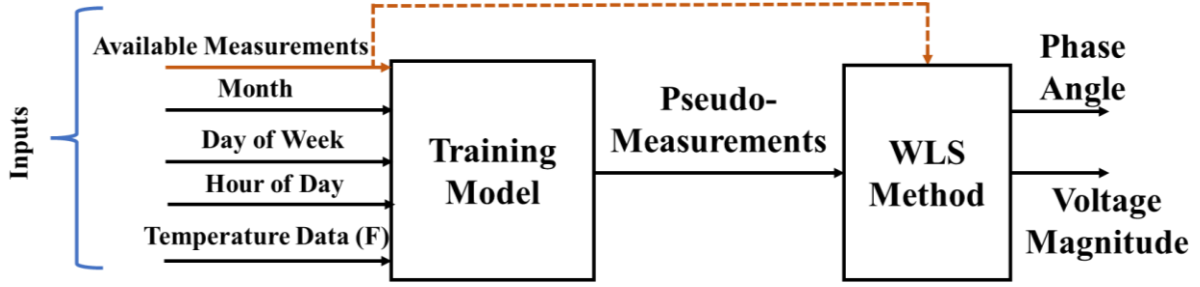


Figure 3- 1. Diagram of the proposed method.

In what follows, the proposed method is described for preparing a training data based on different criteria to generate pseudo-measurements in ADNs where only a limited number of real measurements are available. The overall diagram of the proposed method is shown in Figure 3.2.

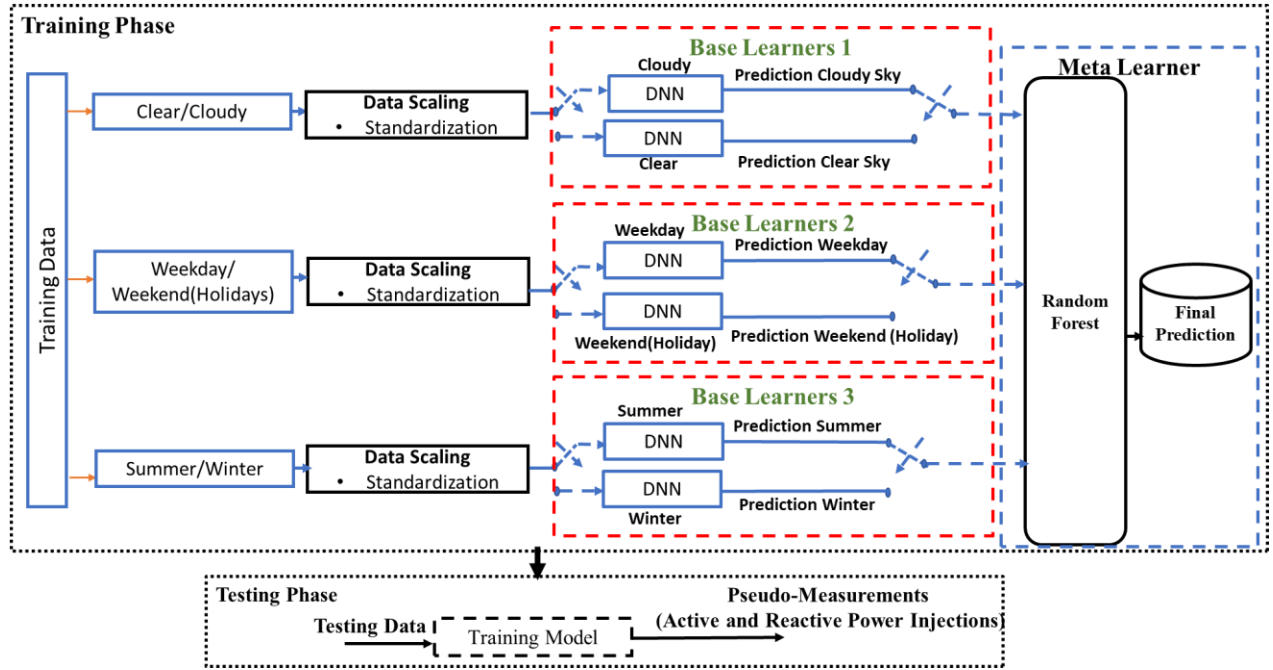


Figure 3- 2. Training and testing diagram of the proposed method.

Training data consists of the available measurements, month of a year, day of a week, hour of a day, and temperature data. Month of a year, day of a week, hour of a day are coded as

shown in Figure 3.3. In order to address different dynamic behaviors in ADNs, the training data are also divided into clear/cloudy, weekday/weekends, and summer/winter training models. These stratifications represent significant differences in power consumption and generation.

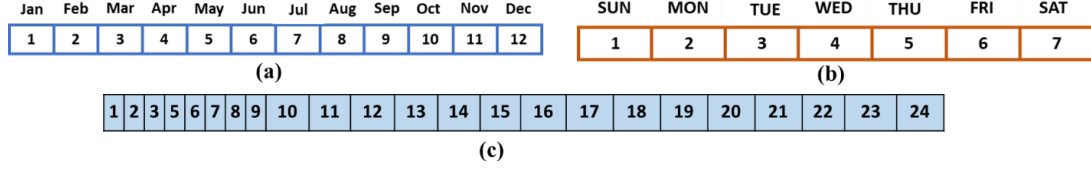


Figure 3- 3. (a) Month of a year, (b) Day of a week, and (c) Hour of a day.

In this study, summer and winter seasons are considered. The actual data for Bozeman, MT [213] are used, as these two seasons are mainly considered for this city. Moreover, the training data are divided based on clear and cloudy skies since photovoltaic panels (PVs) are considered as DGs in this study; outputs of PV depend on solar irradiance [214], which is affected by cloudiness (see Figure 3.4). As can be seen, PV outputs have a severe dynamic behavior in cloudy sky in contrast to a smooth behavior in clear sky. It is worth mentioning that in some cases PV outputs have a combination of smooth and dynamic behaviors. All behaviors of photovoltaic panels (PVs) are considered to improve model robustness and uncertainty in this study.

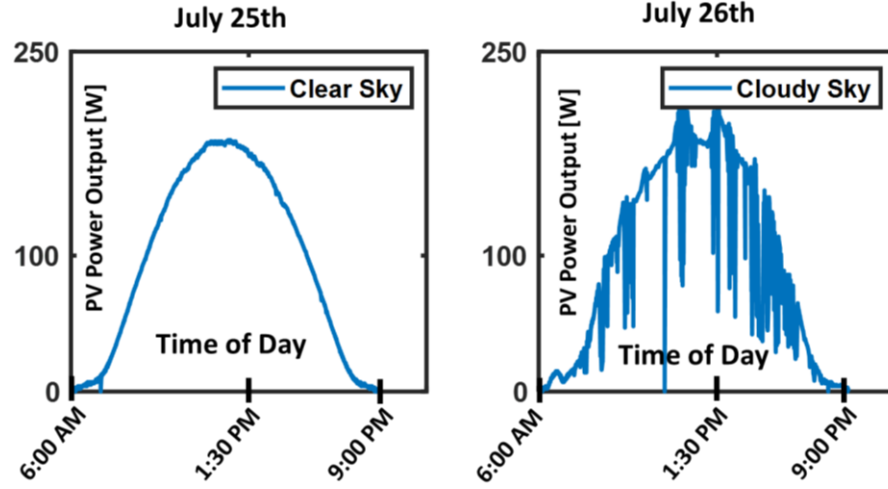


Figure 3- 4. PV outputs over two consecutive days.

After coding month of a year, day of a week, hour of a day and dividing data into clear/cloudy, weekday/weekends, and summer, all data are then split into base learner training, meta learner training, and test sets with ratio of 60%, 20%, and 20%, respectively. After splitting the data, the base learner training data are categorized into three groups (clear or cloudy; weekday or weekend; and winter or summer). Base learner training data for each base learner with different ranges or units are transformed to scaled data by using the standard scaler from Scikit-Learn [215]. Then base learner training data from each category is used to train each base learner which consists of two DNN models. The same hyperparameters are considered for each DNN. These hyperparameters are detailed in Section 3.2.1. Outputs of each base learner are generated based upon the binary inputs category of each base learner (e.g., cloudy, or clear; weekday, or weekend; and winter, or summer). The output values of each base learner are different since different training data are considered for training each base learner. The meta learner training data is then put through the trained base learner system and the results of each base learner are used to train the random forest meta learner.

A random forest is an ensemble of decision trees, and is one of the most popular algorithms for a regression solution because of its simplicity and high accuracy [216]. Since a single tree is not sufficient in providing effective results, the random forest approach has been developed to outperform a decision tree performance. Figure 3.5 illustrates a random forest model that consists of N decision tree models.

Random forest is a tree-based machine learning approach that takes advantage of multiple decision trees for making an accurate decision. In a decision tree, in order to find the output, every node works on a random subset of features. But, in a random forest, the output of multiple individual decision trees is combined to generate the final result. The procedure for combining the several individual models is called Ensemble Learning [214].

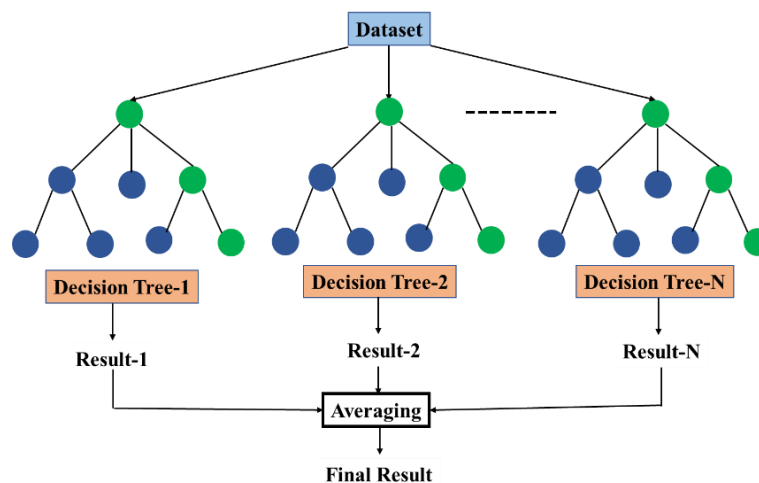


Figure 3- 5. A random forest model.

Also, random forests require less training data than DNNs, so it is anticipated that the performance of the random forest meta learner will outperform each base learner. The result of the meta learner obtained on the unseen test data is presented as the final pseudo-measurement result.

Case Study

The IEEE standard 69 bus distribution network (shown in Figure 3.6) is chosen as the case study for this work and the effectiveness of the proposed method is evaluated using this system. This system is suitably adapted to include mix of commercial and residential loads and DGs. A set of experimental data (available for a time period of one year) obtained from Open Energy Information (OpenEI) [217], is utilized with a time step of 1-hour. Four Photovoltaic (PV) panels are placed [218] on buses 17, 20, 38, and 64 with maximum generation of 22 kW, 11 kW, 17 kW, and 9 kW; respectively. The hourly data of power generation of the photovoltaic system are computed based on the real data from the weather history of Bozeman, MT, USA [219].

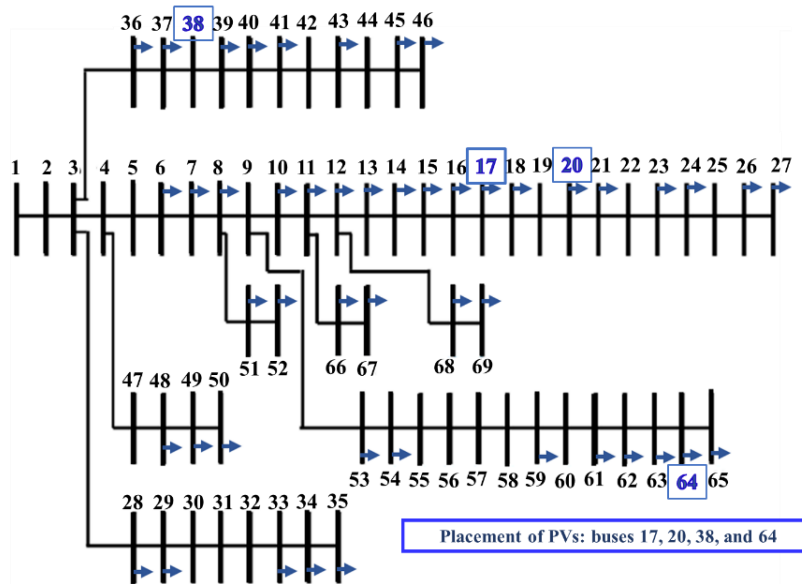


Figure 3- 6. IEEE Standard 69 bus Distribution System.

In this work, it is assumed that two sets of real and reactive power measurements are available (P_{1-2} , Q_{1-2} and P_{40-41} , Q_{40-41}). The inputs and outputs of the proposed method are shown in Table 3.1.

Table 3- 1. .Input and output of the proposed method

Training Data	Output
P_{1-2} , Q_{1-2}	Pseudo-measurements: Buses' Active (P) and reactive (Q) power injections
P_{40-41} , Q_{40-41}	
Month	
Day of week	
Hour of day	
Temperature Data (F)	

Training Phase. The training and testing phases for the case study will be explained in this section. The hyperparameters chosen for a DNN base learner are shown in **Error! Reference source not found. 3.2.**

Table 3- 2. Hyperparameters of each base learner for the proposed and random split methods

Optimizer	Stochastic gradient decent
Number hidden layers	2
Output layer activation	Linear
Hidden layer activation	Tanh
Number of neurons	500 250
Number of outputs	49

For the random forest (meta learner), the hyperparameters of Max Depth = 60 and Number Estimators = 100 are used. All other hyperparameters are set to their default values

using the Scikit-Learn version 0.23.2 package in python and the TensorFlow library 2.4.1 package in python.

Base learner training data is fed into each base learner (base learner₁, base learner₂, and base learner₃) to train each base learner. Then, meta learner training data is fed into a pre-trained base learner and a random forest is trained based on the results of a base learner.

Testing Phase. In this step, the testing data is fed into the training model to evaluate the proposed method. The MPE (Mean Percentage Error) and RMSPE (Root Mean Square Percentage Error) values for each base learner (base learner₁, base learner₂, base learner₃) and the meta learner for the proposed method, and random split considering no DG and 4 DGs are shown in Figure 3.7 (a) and (b), respectively. MPE and RMSPE values calculated from (1) and (2), respectively:

$$MPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{x}_i - x_i}{x_i} \right| \times 100 \quad (1)$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\hat{x}_i - x_i}{x_i} \right)^2} \times 100 \quad (2)$$

The relatively small error values for MPE and RMSPE show that the proposed approach is robust and performs well with or without DGs in DNs when available data is divided into three binary categories based on temporal and meteorological features: clear or cloudy; weekday or weekend; and winter or summer.

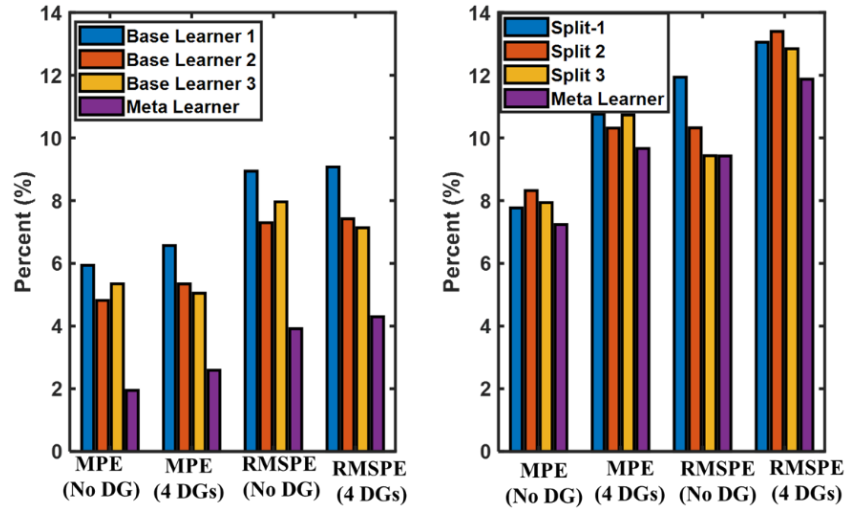


Figure 3- 7. MPE and RMSPE values for (a) the proposed method and (b) and random split method with no DG and four DGs.

To verify the proposed method, pseudo-measurements (active power injection) are generated at two different times. These pseudo-measurements are shown in Figure 3.8 (08/29 at 8:00 AM), and Figure 3.9 (11/04/2015 at 1:00 PM). Moreover, actual and predicted values for bus 38 for 01/31/2015 at 8:00 AM through 02/02/2015 at 5:00 PM is shown in Figure 3.10.

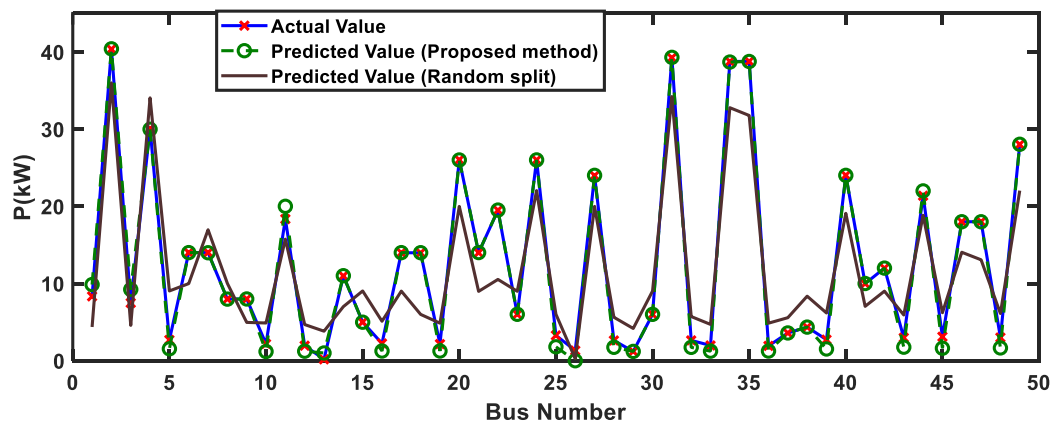


Figure 3- 8. Active power injection at 08/29 at 8:00 AM (Pseudo-Measurements Generated).

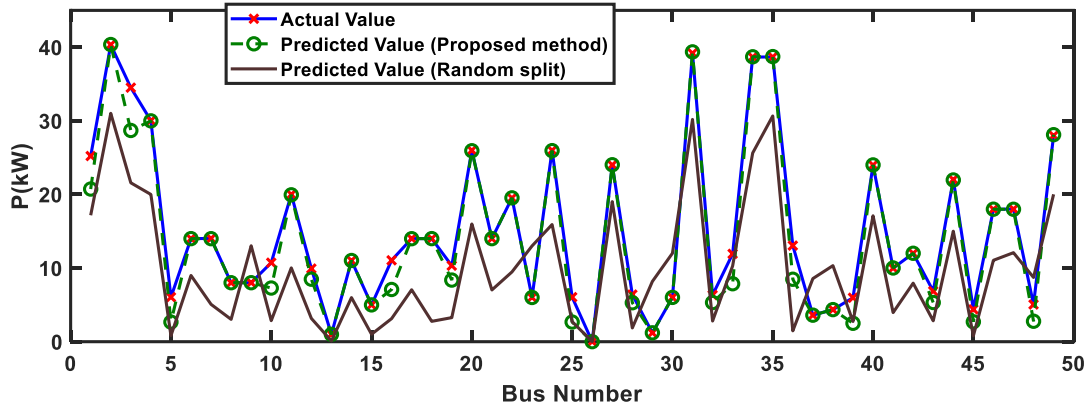


Figure 3- 9. Active power injection at 11/04 at 1:00 PM (Pseudo-Measurements Generated).

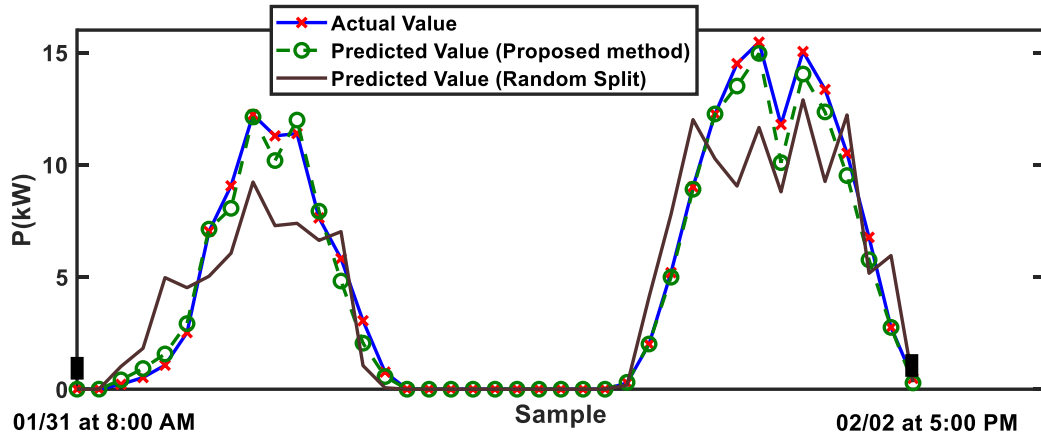


Figure 3- 10. Actual and predicted values for Bus 38.

As it is clear from above figures, the predicted values from the proposed method correctly follow the actual values which are calculated from the power flow calculation compared to split random method.

In order to evaluate the accuracy of the generated pseudo-measurements (active and reactive power injections from the proposed method) real measurements (active and reactive power flows), and network parameters of the case study (IEEE standard 69 bus distribution network) are used in the WLS method to calculate state variables of a DN. The state vector is

defined as $\hat{x} = [V_1, \dots, V_N, \delta_1, \dots, \delta_N]$, where V, δ are the estimated values of voltage magnitude and voltage angle at each bus, respectively, and N is the number of buses. Bus 1 is considered as a reference bus with $\delta_1 = 0$. The standard deviation is considered as 50% of the nominal value for pseudo-measurements and 3% of the actual value of active and reactive power flow measurements. To verify the accuracy of the WLS method, root mean square error (RMSE) for voltage magnitude estimations of the proposed and the random split approaches are shown in Table 3.3.

Table 3- 3. RMSE (%) values for voltage magnitude estimation for the proposed method and random split

	Proposed Method	Random Split
RMSE (%)	0.080	0.105

As it is clear, the result of voltage magnitude estimations is improved when pseudo-measurements generated from the proposed method are used along with real measurements using a WLS method.

Conclusion

In this chapter, a data-based approach is proposed to generate pseudo-measurements in ADNs by considering limited availability of real measurements along with temporal and meteorological information. Two sets of active and reactive power flow measurements, month of year, day of week, hour of day, and the temperature data, are considered as inputs for generating pseudo-measurements.

In order to train an effective model to be robust against dynamic behaviors of ADNs, the training data are divided based on temporal and meteorological features including season

patterns (summer and winter), weekdays and weekends, cloudy and clear sky data. Three base learners are considered for each category. Each base learner consists of two DNNs: one network for one part of the binary category (e.g., cloudy), and one network for each of the other parts of the category (e.g., clear).

In total, six DNNs with identical hyperparameters are considered as base learners. Outputs of these three base learners are generated based upon the binary inputs category of each base learner (e.g., cloudy or clear). The predictions from each base learner are considered as inputs for training a random forest meta learner to calculate the pseudo-measurements results.

In this chapter, it is shown that a DNN-only approach cannot work well when there is not enough data for training a model. This problem is addressed by using a DNN approach along with another method to address a lack of measurement to generate accurate pseudo-measurements in ADNs.

The results show that the proposed method is accurate enough in the absence and presence of DGs in different criteria compared with random split method. In order to evaluate the accuracy of pseudo-measurements for the DSSE calculation, the WLS method is performed by using real measurements and pseudo-measurements generated from the proposed method.

The results of the WLS method show that pseudo-measurements are sufficiently accurate to perform WLS and obtain state estimation values to make the network observable.

CHAPTER FOUR

DISTRIBUTION SYSTEM STATE ESTIMATION AND FALSE DATA INJECTION ATTACK
DETECTION WITH A MULTI-OUTPUT DEEP NEURAL NETWORK

Distribution System State Estimation (DSSE) has been introduced to monitor distribution Networks (DNs); however, because of the transition from passive to active distribution networks, traditional DSSE methods are not able to reveal the operational conditions of Active Distribution Networks (ADNs). Furthermore, DSSE calculation depends heavily on real measurements from measurement devices in DNs. Therefore, smart measurement devices are being gradually installed in DNs to enable more reliable performance and obtain accurate DSSE calculations from real measurements.

However, despite the progress toward the development and deployment of advanced measuring instruments in DNs, attackers attempt to disrupt the performance of DNs by manipulating the data obtained from measurement devices. A malicious attacker's goal is to falsify the real measurements provided by measuring tools. The accuracy of DSSE results can be significantly affected by such False Data Injection Attacks (FDIAs). If FDIAs cannot be correctly identified and detected by an appropriate method, the control center performance will be significantly degraded and diminished. DSSE results are used by a control center for further processing including optimal power flow, contingency, voltage stability, and network security analysis, etc., so poor DSSE results can have far-reaching negative consequences.

The concept of False Data Injection Attacks (FDIAs) in power systems has been studied and applied to allow Bad Data Detection (BDD) on measurement information. Traditional BDD schemes are performed using the ℓ_2 -norm of the measurement residual after SE calculations

[220], [221]. However, FDIAs can successfully bypass most conventional BDD algorithms and attackers can inject bad data into measurement data and introduce arbitrary errors into the output of the SE [222].

Unlike in DNs, SE techniques, attacks, and defense approaches are well developed in transmission networks (TNs). However, due to the differences between TNs and DNs, these techniques cannot be directly applied in DNs. Moreover, renewable energy sources and advanced technologies are becoming more common in DNs [223]–[225], and the security aspects of ADNs—including the ability to detect cyberattacks such as FDIAs [226], [227]—are required to improve ADN performance. Unfortunately, conventional FDIA detection techniques are often unable to identify FDIAs into measurement data in DNs.

In the last decade, machine learning approaches have been widely used and developed for control and monitoring in power networks [228]–[231]. In [232], a machine learning approach is used for an energy storage program and load management in power networks. Moreover, due to limitations in detecting FDIAs using conventional (model-based) BDD methods, machine learning approaches have been extensively applied to identify malicious manipulation of measurement data. Faster performance and accurate results will benefit from machine learning technologies over conventional BDD methods [233]. Machine learning algorithms are based on the data obtained from the power networks, unlike model-based approaches which are based on network topologies as well as measurement data.

In this chapter, the focus is on identifying FDIAs on measurement data using a DNN approach in DNs. FDIA detection is considered a supervised binary classification problem. Some existing research for detecting FDIAs using ML approaches is discussed in 2.7 section.

In the existing research, DSSE calculation and bad data detection on state variables and measurement data are performed separately by using model-based or data-based approaches. In some cases, a single DNN model is designed to solve multi-output classification problems. In contrast, a single DNN model is introduced to simultaneously perform DSSE and FDIA detection on PMU measurements. Different scenarios are considered for FDIAs to verify the proposed method; moreover, the performance of the proposed method is compared to when DSSE calculation and FDIA detection are performed by separate DNN models [234]. To make a comparison between the robustness of data-based and model-based approaches, the WLS method is performed to obtain state variables in the presence of FDIAs on PMU measurements.

The main contributions of this work are summarized as follows:

- 1- A single DNN model is designed to perform the DSSE calculation and FDIA detection on PMU measurements simultaneously. The results are compared to when DSSE calculation and FDIAs are performed separately using two independent DNN models.
- 2- Having $N + 1$ classification nodes, where N is the number of PMU measurements, allows the DNN to identify which PMU measurement, if any, was affected by FDIA, or if none of the measurements were affected.
- 3- The performance of the proposed method is investigated for FDIA detection on PMU measurements with different attack scenarios.
- 4- To make a comparison between data-based and model-based approaches, DSSE calculations are performed using the WLS as a model-based approach.
- 5- The effectiveness of the proposed method is tested for passive and active distribution networks: the 33 and 69 IEEE distribution networks, respectively.

- 6- It is shown that the proposed method accurately calculates state estimation variables, even in the presence of erroneous measurements.
- 7- The execution time comparison among the proposed method, disjoint DNN model for DSSE calculation and FDIA detection, and model-based approach are calculated. The results show that the average execution time is decreased significantly when DSSE calculation and FDIA detection on measurements are done simultaneously.

Power System State Estimation

State estimation calculations are essential for continually improving the performance and management of power networks [235]. Distribution System State Estimation (DSSE) aids to enhance monitoring and controlling of DNs effectively and efficiently. Moreover, state estimation results can be utilized for load forecasting, stability analysis, optimal power flow, bad data detection, and energy market [236]. The Weighted Least Square (WLS) method is one of the traditional methods which is effective in both distribution and transmission networks.

WLS formulation for State Estimation

The general measurement model for the state estimation problem can be expressed as:

$$z = h(x) + e_z, \quad (1)$$

where $z = [z_1 \dots z_M]^T$ is the vector of the measurements, and M is the number of available measurements; $h = [h_1 \dots h_M]^T$ is the list of measurement function vectors, and it is commonly nonlinear. The relationship between the available measurements and state vectors are shown by $h(x)$. $x = [x_1 \dots x_N]^T$ is the state variable vector and N is the number of state variables. Lastly, $e_z \sim \mathcal{N}(0, R_z)$ is the measurement noise vector, and it is assumed to be zero mean and Gaussian

random variable with covariance matrix $R_z = \text{diag}\{\sigma_{z_1}^2, \sigma_{z_2}^2, \dots, \sigma_{z_m}^2\}$. For instance, power flow equations can be expressed as:

$$P_i = V_i \sum_{j=0}^N V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (2)$$

$$Q_i = V_i \sum_{j=0}^N V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}). \quad (3)$$

In these equations, P_i and Q_i are the real and reactive power injections at bus i , respectively. G_{ij} and B_{ij} are the real and imaginary part of the nodal admittance matrix element Y_{ij} and $\theta_{ij} = \theta_i - \theta_j$ is the standing phase angle difference between buses i and j .

Additional equations describe the active (P_{ij}) and reactive (Q_{ij}) power flow from bus i to bus j :

$$P_{ij} = V_i V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) - G_{ij} V_i^2 \quad (4)$$

$$Q_{ij} = V_i V_j (G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij}) + B_{ij} V_i^2 \quad (5)$$

When WLS is performed for the state estimation calculation, the objective function is defined as:

$$J(x) = \sum_{i=1}^M w_i [z_i - h_i(x)]^2 \quad (6)$$

In this formulation, w_i is the weight associated with the i^{th} measurement, and M is the number of available measurements to perform SE calculation. The difference between measurement values (z) and the function corresponding to the state vector ($h(x)$) is expressed as $z - h(x)$, and it is called the measurement residual in the literature. Equation (6) can be defined in matrix form as:

$$J(x) = [z - h(x)]^T W [z - h(x)] \quad (7)$$

$W_{M \times M}$ is a diagonal matrix, whose diagonal elements correspond to the weights w_i . The Iterative Gauss-Newton (IGN) method is commonly performed to minimize the objective function $J(x)$; [237]. In IGN, the following function is solved at each iteration k :

$$G(x_k)\Delta x_k = H_k^T W [z - h(x_k)] \quad (8)$$

where $H_k = H(x_k)$ is the Jacobian matrix at iteration k , and $G(x_k) = H_k^T W H_k$ is the Gain matrix. Δx_k is the updated state vector to calculate the new state as the following:

$$x_{k+1} = x_k + \Delta x_k \quad (9)$$

The iterative calculation continues until a predefined convergence criterion is reached. The largest absolute value of the updated state vector (Δx_k) is compared to a predefined tolerance threshold ε . When $\max(|\Delta x|) < \varepsilon$, the calculation will be stopped. The state vector will be estimated in the last iteration by the WLS approach.

The state vector of the power grid can be defined as a set of variables; when the state variables are calculated, other electric power quantities could be computed from these states [164], [165]. In Node Voltage Distribution System State Estimation (NV-DSSE), voltage magnitudes and phase angles for all buses are considered as state variables. State vectors can be represented in polar coordination as $x = [\delta_2, \dots, \delta_N, V_1, \dots, V_N]$, where δ_N , V_N are voltage phase angle and magnitude, respectively, and N is the number of buses. It is assumed that there are no measurement devices installed in the slack bus and only conventional measurements are available in the DN. The phase angle of the slack bus is zero ($\delta_1 = 0$). However, if there is a measuring device at the slack bus, the state vector will be defined as $x = [\delta_{1\emptyset}, \dots, \delta_{N\emptyset}, V_{1\emptyset}, \dots, V_{N\emptyset}]$ where the phase angle $\delta_{1\emptyset}$ is not zero any more [238]–[240].

There are two main differences between the non-PMU configuration and the PMU configuration. First, the definition of the mathematical equation relating measurements to physical parameters of the DNs is altered. Second, the Jacobian matrix has a different structure. [241].

False Data Injection Attacks (FDIAs)

State estimation is needed to evaluate the operating status of power networks, including voltage magnitude and phase angle of buses from available measurements. Unfortunately, the state variables can be manipulated by injecting FDIAs into meter measurements [57], which reduces the accuracy of DSSE results.

The measurement vector z could be manipulated and changed to a falsified measurement z_a when attackers inject malicious data:

$$z_a = z + a. \quad (10)$$

In this formulation, $a \in \mathbb{R}^{m \times 1}$ is the malicious data vector that is injected into the measurement vector. The erroneous measurement vector, z_a , can lead to an inaccurate system state $\hat{x}_a = x + c$, where c is the resultant error in the state vector. The FDIAs cannot be identified by bad data detection approaches if an attacker knows the structure of the power system h . For example, the FDIAs can bypass the BDD if $a = h(\hat{x}_a) - h(\hat{x})$, which causes the residual error before and after the attack to be the same:

$$\begin{aligned} r_a &= \|z_a - h(\hat{x}_a)\| \\ &= \|z + a - h(\hat{x}_a) + h(\hat{x}) - h(\hat{x})\| \\ &= \|z - h(\hat{x}) + a - h(\hat{x}_a) + h(\hat{x})\| = r \end{aligned} \quad (11)$$

The general architecture of FDIAs on measurements and DSSE procedures are shown in Figure 4.1. The measurement vector is manipulated by the FDIA vector (a), which modifies it to become $z_a = z + a$. Falsified measurements and network topology are fed into DSSE calculation which can be performed using model-based or data-based approaches. Then, the SE results are sent to the control and management center for further processing, including bad data detection using appropriate methods.

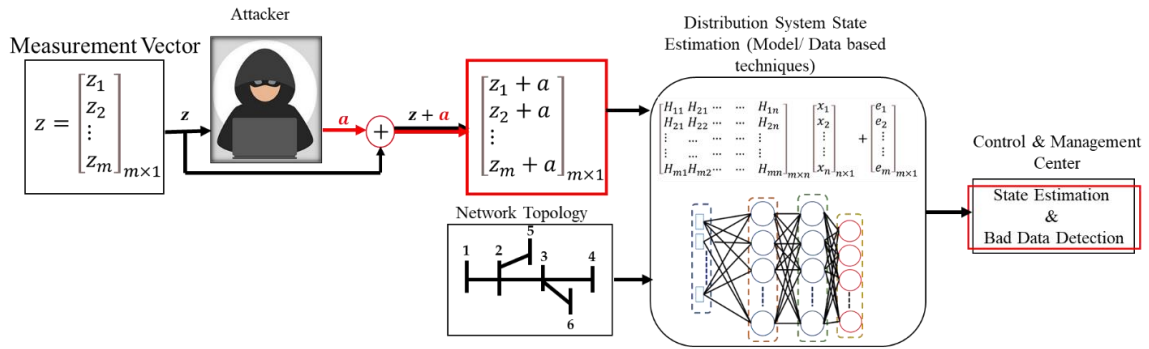


Figure 4- 1. General architecture of FDIAs on measurements and DSEE procedure.

Machine Learning Approach to Detect FDIAs

Machine learning is a form of artificial intelligence that gives computers an ability to learn without being explicitly programmed [242]. FDIA detection is defined as a supervised binary classification problem. The main objective of a binary classifier for FDIAs is to classify measurements as being either secure (z) or attacked ($z_a = z + a$). A binary classification problem can be defined as:

$$y_i = \begin{cases} 0, & \text{if } a = 0 \\ 1, & \text{if } a \neq 0 \end{cases} \quad (12)$$

where $y_i = 0$ and $y_i = 1$ indicate there is no attack or there is an attack on a measurement, respectively, and a is the attack vector.

The Deep Neural Network (DNN) is a subset of machine learning inspired by the organization or structure of the human brain. DNN is one of the fastest growing artificial intelligence technologies. DNN methods have been proposed widely to detect FDIAs with high accuracy [243]; however, this technique requires more time and data for a training phase [244]. In feed-forward DNN models, the information flows in only one single direction from the input, through optional hidden layers, to the output, as shown in Figure 4.2.

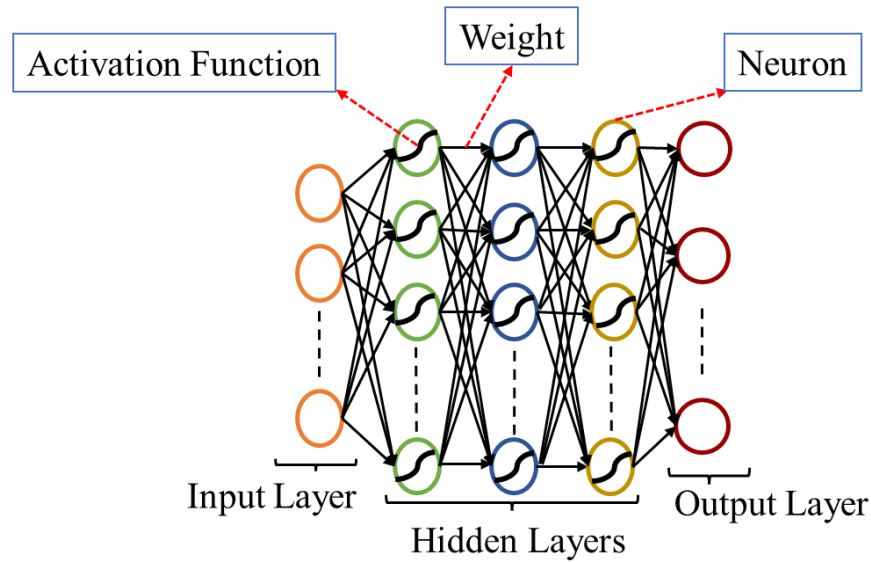


Figure 4- 2. DNN configuration

A DNN consists of activation functions, weights, neurons, an input layer, hidden layers, and an output layer. The input layer comprises neurons that receive the input variables and transfer them to subsequent layers in the network. The number of neurons in the input layer must be the same as the number of the features or attributes in the dataset. The hidden layers are placed between the input and the output layers; the number of hidden layers and the number of

neurons in each layer are determined experimentally. The weights in the network are constantly updated so the output can reliably predict an outcome based on the original input. The strength or the magnitude of connection between two neurons is called a weight. The value of the weights is usually small and falls within the range of 0 to 1. Neurons have two important roles: first, they determine the sum of the weighted inputs, and second, they initiate an activation process to normalize the sum. Weights are associated with each input of the neuron. The network learns these weights during the learning phase. The activation function, which can be either linear or nonlinear, is the decision-making center at the neuron output. Three common activation functions are sigmoid, tanh, and Rectified Linear Unit (ReLU).

Methodology

As mentioned earlier, DSSE calculation and FDIA detection are typically performed on measurements separately. In conventional cases, DSSE calculation is performed in the first stage, and then BDD is performed to identify FDIAs on measurements from SE results.

In data-based approaches, as shown in Figure 4.3 (a) and (b), two separate DNNs are considered: one to execute DSSE calculations and one to perform the binary FDIA detection. In this study, as shown in Figure 3 (c), a single DNN model is designed to perform DSSE calculation and FDIA detection simultaneously. The proposed method is compared when DSSE calculation and BDD on measurements are performed by using two independent DNN models. Regression-based and classification-based DNNs are applied to perform DSSE calculation and detect secure and attacked measurements, respectively. In the regression problem, a training model tries to learn the relationship between inputs (measurements: z) and outputs (state variables: x) in a training step ($z \rightarrow x$). Then, a learning model is used in the testing phase to

calculate the corresponding outputs from a testing data. For binary classification, a training model is used to classify the measurements as either secure or compromised. It is assumed the attacker injects false data into the original measurements by directly modifying the measurement vector: $z \rightarrow z_a$. (This can also be modeled as $z_a = z + a$.)

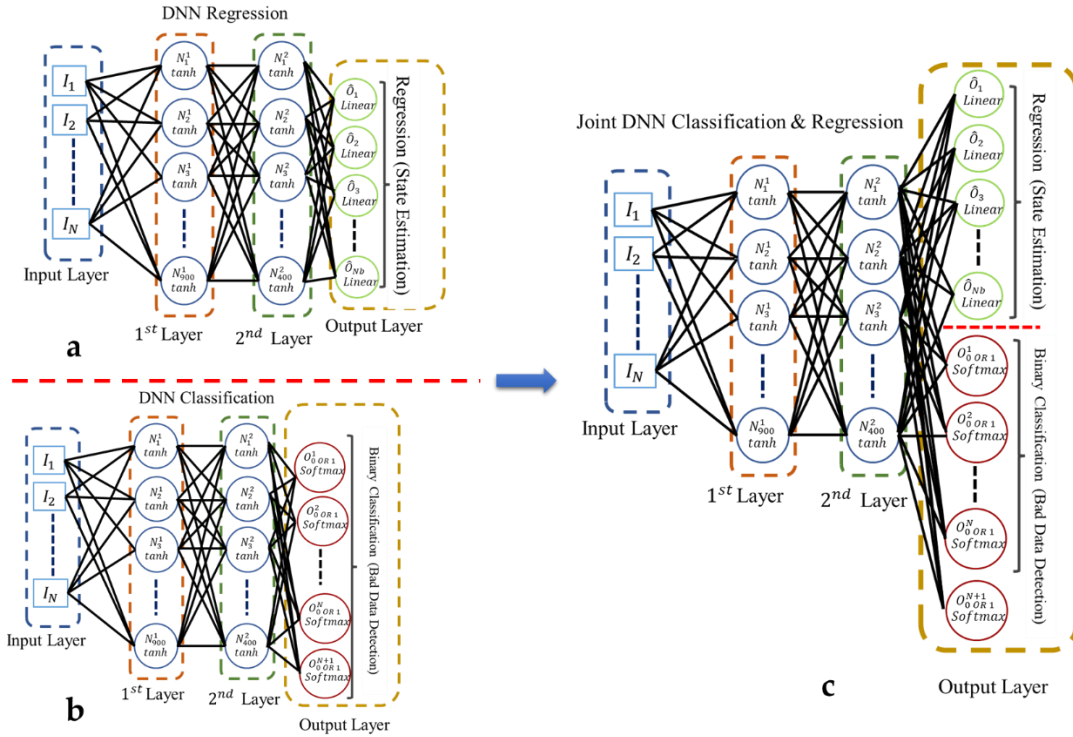


Figure 4- 3. (a) A regression DNN architecture for DSSE calculation; (b) A binary classification DNN model for FDIAs detection on measurements; (c) the proposed DNN model configuration to perform DSSE calculation and FDIAs on measurements simultaneously.

In this study, the bus voltage and branch current magnitudes are considered as available measurements which are obtained from the PMU devices installed on a limited number of buses in the distribution network.

The attack models (V_a and I_b) on voltage and branch current magnitudes are expressed in (13) and (14):

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \dots \\ V_N \end{bmatrix} \rightarrow V_a = \begin{bmatrix} V_{a,1} \\ V_{a,2} \\ \dots \\ V_{a,N} \end{bmatrix} \quad (13)$$

$$I = \begin{bmatrix} I_1 \\ I_2 \\ \dots \\ I_N \end{bmatrix} \rightarrow I_b = \begin{bmatrix} I_{b,1} \\ I_{b,2} \\ \dots \\ I_{b,N} \end{bmatrix} \quad (14)$$

The proposed DNN model consists of input, two hidden, and output layers. The number of neurons in the input layer is equal to number of available PMU measurements (N). The first and second hidden layers have 900 and 400 neurons, respectively. $\tanh$ function is considered as the activation function for both hidden layers. The output layer consists of N_b regression nodes and $N + 1$ classification nodes. N_b is the number of state variables: the voltage magnitudes and phase angles of all buses. N is the number of PMU measurements in the distribution networks: one voltage magnitude and one current magnitude per PMU. The output nodes corresponding to regression use a linear activation function. The classification nodes in the output layer use a SoftMax activation function, allowing the algorithm to identify only one output node as being the most likely output. Having $N + 1$ classification nodes allow the DNN to identify which PMU measurement, if any, was affected by FDIA. Note that if multiple measurements are attacked with the same vector injection, the algorithm will only report one of them because the softMax function is used as the classification layer activation function. Stochastic Gradient Decent (SGD) is used as the optimizer. All other hyperparameters are set to their default values using the TensorFlow library 2.4.1 package in python.

In order to evaluate the DSSE results (i.e., the regression outputs), Mean Percent Error (MPE) and Mean Absolute Error (MAE) are calculated using (15) and (16):

$$MPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{x}_i - x_i}{x_i} \right| \times 100 \quad (15)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{x}_i - x_i| \times 100. \quad (16)$$

In both equations, $\hat{x}$ is the estimated value, x is the actual value, and n refers to the data set size.

Accuracy and F₁- score, defined below, are used for binary classification outputs:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (17)$$

$$F_1 = \frac{2 \times precision \times recall}{precision + recall}. \quad (18)$$

In (17), TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives. Precision and recall are calculated as follows:

$$Precision = \frac{TP}{TP+FP} \quad (19)$$

$$Recall = \frac{TP}{TP+TN}. \quad (20)$$

A confusion matrix can be used to analysis the performance of a given classifier.

Correctly classified and misclassified outcomes are represented on the on and off diagonals of the confusion matrix, respectively. Where these values are nonzero, the entries are highlighted as blue (correct) or red (incorrect).

Results

Case Study I

The IEEE 33 bus distribution network (shown in Figure 4.4) is chosen as one of case studies and the effectiveness of the proposed method is evaluated using this system. “True” values of electrical parameters are calculated by power flow analysis.

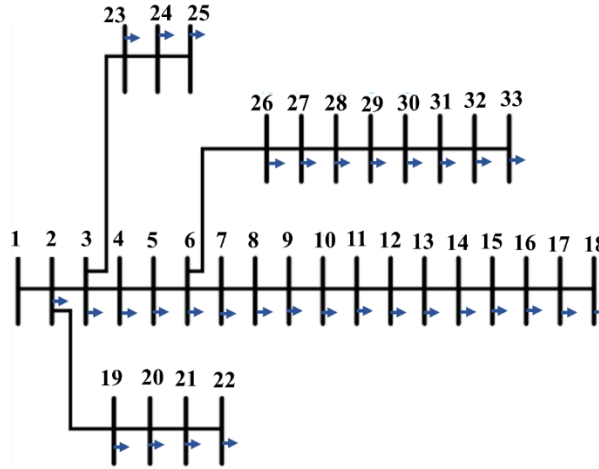


Figure 4- 4. IEEE standard 33 bus distribution network.

Measurements are randomly obtained according to their probability density function for each Monte Carlo trial. The following assumptions are considered:

- 1- The number of Monte Carlo trials is chosen as $N_{MC} = 12000$.
- 2- A Gaussian distribution, with $3\sigma = 50\%$ of the nominal value, is considered for power injection on the buses.
- 3- Three PMUs [245]—located at buses 9, 16, and 31—are assumed as the available measurement devices in the network. Voltage and branch current magnitudes from each

PMU are used for FDIA detection and performing DSSE using the proposed method. It is assumed 12000 samples for each PMU according to the number of Monte Carlo trials.

- 4- Two types of attack vectors are injected into available measurements:
 - a) $z_a = \{(1 \pm .05)z, (1 \pm 0.1)z\}$, as in [201].
 - b) $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$, where μ_z and σ_z are the average and standard deviation of each measurement vector, respectively. In this case, it is assumed that the attacker knows the distribution of each measurement and wants to falsify measurements based on the true measurement distribution.
- 5- Data are then split into train and test sets with the ratio of 67% and 33%.
- 6- The pseudo-measurements of active and reactive power injections and flows are generated to make the system observable and perform WLS calculations with the inclusion of the PMU measurements.
- 7- Voltage magnitudes and phase angles for all the buses are considered as a state variable: $x = [\delta_2, \dots, \delta_N, V_1, \dots, V_N]$, where δ_N , V_N are the voltage phase angle and magnitude, respectively, and N is the number of buses. It is assumed that there are no measurement devices installed in the slack bus and $\delta_1 = 0$.
- 8- The standard deviation is considered as 50% of the nominal value for pseudo-measurements and 3% of the actual value of active and reactive power flow measurements. A gaussian error, with $3\sigma = 1\%$, is added to PMU measurements (voltage and branch current magnitudes) to model uncertainty. To model the zero injection buses, virtual measurements' error is considered as 10^{-8} .

False Data Injection Attacks on Measurements with $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$. In this case, FDIAs are injected to measurement samples randomly, and 10% of each available measurement ($N_a = 1200$) are falsified randomly by FDIAs ($z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$).

In Figure 4.5, true and attacked measurements for 120 consecutive samples are shown. As it is clear, after FDIA, the true measurements are changed to falsified values, and these erroneous measurements are identified using the proposed method.

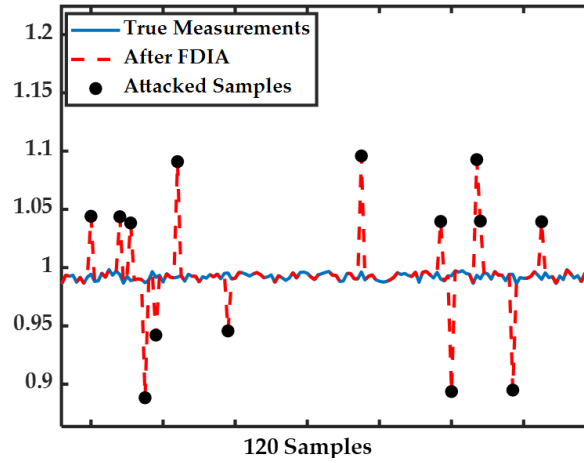


Figure 4- 5. Original and attacked measurements for 120 samples when $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

Next, the FDIA detection and DSSE results from the proposed method are discussed. Also, the appropriate binary classification-only or regression-only DNN model and the WLS method are explained. All model performances are analyzed for the case where 10% of each available measurement $N_a=1200$ are falsified randomly by FDIAs ($z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$).

Table 4.1 shows the results for the proposed and appropriate binary classification-only methods. It shows that the proposed method successfully detected the FDIAs on the PMU

measurements when false data $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$ are injected randomly in 10% of available measurements. The accuracy and F1-score of the proposed method are 0.930 and 0.889, respectively, which are better than when performing the binary classification only (i.e., 0.882 and 0.814, respectively).

Table 4- 1. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification- only DNN model when $N_a = 1200$ (10%), and $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

		Joint DNN Regression and Classification							DNN Classification Separately						
Actual Values	FDIA	$ V_9 $	421	3	0	0	0	0	0	0	0	0	0	179	0
		$ V_{16} $	110	260	0	0	0	0	0	0	0	0	122	0	0
		$ V_{31} $	71	0	229	3	7	4	0	0	228	3	3	80	0
		$ I_{9-10} $	40	0	0	266	2	2	0	0	0	268	0	42	0
		$ I_{16-17} $	36	0	0	0	201	0	0	0	0	0	199	38	0
		$ I_{31-32} $	1	0	0	0	0	228	0	0	0	0	0	229	0
	No FDIA		0	0	0	0	0	0	0	0	0	0	0	0	2076
		$ V_9 $	$ V_{16} $	$ V_{31} $	$ I_{9-10} $	$ I_{16-17} $	$ I_{31-32} $	No FDIA							
		Predicted Values							Predicted Values						
		Accuracy = 0.9295 F_1 -Score = 0.8885							Accuracy= 0.8820 F_1 -Score = 0.8142						

The results of DSSE calculation from the proposed, the regression-only, and WLS methods are shown in Figure 4.6. As it is clear from this figure, the proposed method has a better performance in both MPE and MAE compared to the other methods in the presence of bad data. When no bad data is present, both of the DNN-based methods (regression-only and combined regression and classification) have similar performance for the state estimation. The WLS estimator has the worst performance in both cases, and the performance is significantly worse in the presence of bad data measurements.

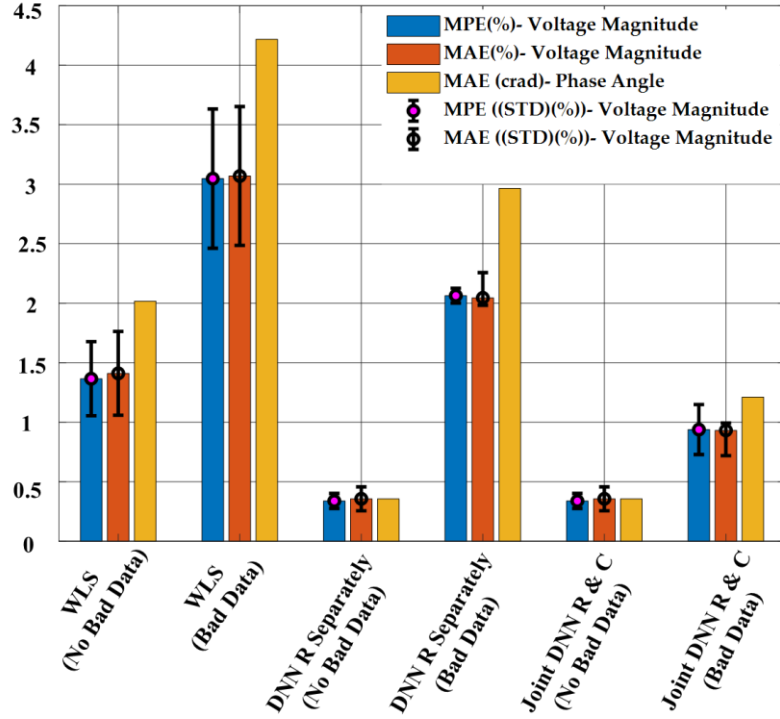


Figure 4- 6. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 1200$ (10%), and $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

False Data Injection Attacks on Measurements with $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$. In this case, it is assumed that the attacker knows the distribution of each measurement and wants to falsify measurements based on the true measurement distribution.

In Figure 4.7, 103 measurements before and after FDIAs are shown. Falsified measurements are identified by applying the proposed method. The FDIA detection and DSSE regression results are shown when FDIAs vector is constructed with $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$. It is clear from the figure that attacked samples come from the same distribution as true measurement samples, making them more difficult to identify.

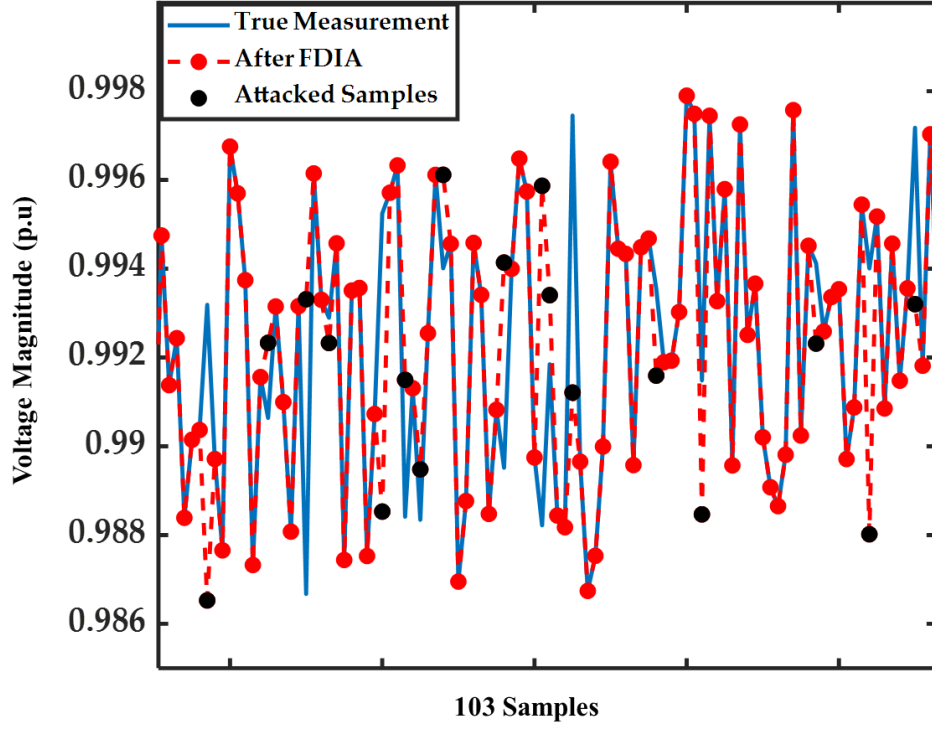


Figure 4- 7. Original and attacked measurements for 103 samples when $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

The results in Table 4.2 for the proposed and appropriate binary classification-only methods show that the proposed method successfully detected most FDIAs on all three PMUs when false data $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$ were injected randomly to 5% of the available measurements. Furthermore, it shows that the proposed method works better than an independent binary classification method. The accuracy and F1-score of the proposed method (0.924 and 0.556, respectively) are better than when performing binary classification only (0.909 and 0.403, respectively).

Table 4- 2. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification-only DNN model when $N_a = 600$ (5%), $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

			Joint DNN Regression and Classification							DNN Classification Separately						
Actual Values	FDIA	$ V_9 $	74	8	1	1	2	1	7	45	0	1	0	0	0	48
		$ V_{16} $	6	74	0	0	0	0	3	5	44	0	0	0	0	34
		$ V_{31} $	0	1	57	0	0	0	18	7	0	29	0	0	0	40
		$ I_{9-10} $	2	0	0	16	2	0	57	0	0	0	1	0	0	76
		$ I_{16-17} $	0	0	1	1	1	1	71	0	0	0	0	0	0	75
		$ I_{31-32} $	0	0	0	0	0	1	73	0	0	0	0	0	0	74
	No FDIA	0	0	4	13	7	20	3437	0	0	0	0	0	0	0	3481
			$ V_9 $	$ V_{16} $	$ V_{31} $	$ I_{9-10} $	$ I_{16-17} $	$ I_{31-32} $	No FDIA	$ V_9 $	$ V_{16} $	$ V_{31} $	$ I_{9-10} $	$ I_{16-17} $	$ I_{31-32} $	No FDIA
		Predicted Values							Predicted Values							
		Accuracy = 0.9242 F_1 -Score = 0.5556							Accuracy= 0.9090 F_1 -Score = 0.4028							

Figure 4.8 shows the results of DSSE calculation from the proposed method, the regression-only method, and the WLS method.

It is clear from the figure that the proposed method has a better performance in the MPE and MAE criteria compared to other methods in the presence of bad data. Similar to what was seen in the previous case study, when no bad data is present, both DNN-based methods (regression-only and combined regression and classification) have similar performance for state estimation. The WLS estimator has the worst performance in both cases, and the performance is significantly worse in the presence of bad data measurements.

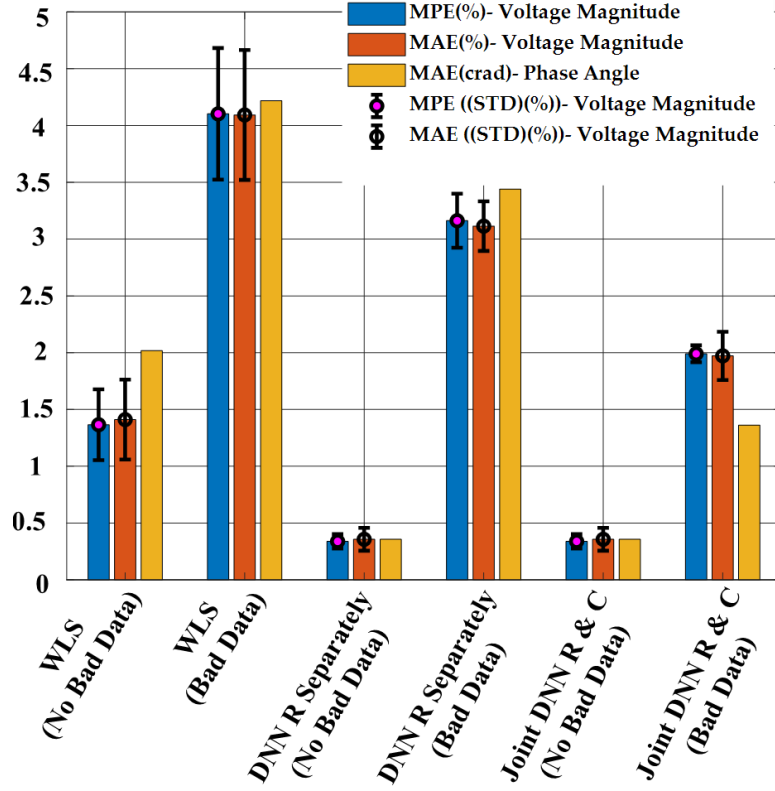


Figure 4- 8. DSSE Results Obtained Using the Proposed Method, the Regression- only DNN Model, and WLS with or without Bad Data when $N_a = 600$ (5%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Table 4.3 shows the results for the proposed method and the appropriate binary classification-only method, showing that the proposed method successfully detected most FDIAs on all three PMUs when false data $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$ were injected randomly to 10% of the available measurements. It also shows that the proposed method works better than an independent binary classification method. The accuracy and F1-score of the proposed method (0.907 and 0.856, respectively) are better than when performing binary classification only (0.839 and 0.403, respectively).

Table 4- 3. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification-only DNN model when $N_a = 1200$ (10%), $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

			Joint DNN Regression and Classification							DNN Classification Separately						
Actual Values	FDIA	$ V_9 $	271	25	39	22	10	28	5	91	4	3	0	0	0	68
		$ V_{16} $	0	312	11	5	5	3	3	13	73	1	0	0	0	60
		$ V_{31} $	2	1	283	8	1	15	4	13	0	65	0	0	0	71
		$ I_{9-10} $	0	0	0	244	31	22	0	4	0	0	2	0	0	143
		$ I_{16-17} $	0	0	0	0	235	22	5	0	0	0	0	0	0	116
		$ I_{31-32} $	3	0	0	0	0	222	5	0	0	0	0	0	0	140
	No FDIA	81	0	0	0	13	0	2024	0	0	0	0	0	0	0	3093
			$ V_9 $	$ V_{16} $	$ V_{31} $	$ I_{9-10} $	$ I_{16-17} $	$ I_{31-32} $	No FDIA	$ V_9 $	$ V_{16} $	$ V_{31} $	$ I_{9-10} $	$ I_{16-17} $	$ I_{31-32} $	No FDIA
		Predicted Values							Predicted Values							
		Accuracy = 0.9068 F_1 -Score = 0.8557							Accuracy= 0.8393 F_1 -Score = 0.4028							

Figure 4.9 shows the results of DSSE calculation from the proposed method, the regression-only, and the WLS methods. It is clear from the figure that the proposed method has a better performance in MPE and MAE criteria compared to other methods in the presence of bad data. Once again, when no bad data is present, both the DNN-based methods (regression-only and combined regression and classification) have similar performance for state estimation. The WLS estimator has the worst performance in both cases, and the performance is significantly worse in the presence of bad data measurements.

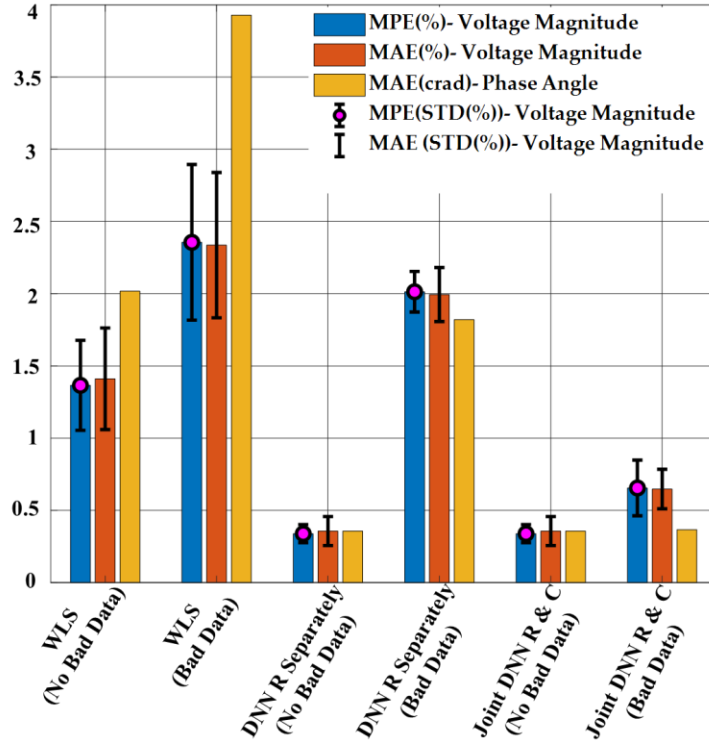


Figure 4- 9. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 1200$ (10%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Table 4.4 shows the execution time for the proposed regression-only method and the WLS methods when the testing data are the same for each method. As it is clear, the execution time is decreased significantly when regression and binary classification are performed simultaneously by a joint DNN model. The WLS execution time is ten times the other methods because, as mentioned earlier, the Jacobian matrix, which is based on physical parameters of a network is recalculated in each iteration, which increases the execution time. Therefore, by applying the proposed method, which is a data-based approach, the execution time is decreased significantly.

Table 4- 4. Execution time for the proposed, regression-only, and WLS methods.

	Joint DNN Regression and Classification		DNN Regression Separately		WLS	
FDIA	$ V $	θ	$ V $	θ	$ V $	θ
$a = [\pm 90\%, 105\%]$	0.62 (s)	0.42 (s)	1.35 (s)	0.78 (s)	30.91 (s)	
$a = [mean, STD]$ $N = 5\%$	0.40 (s)	0.47 (s)	0.35 (s)	0.58 (s)	23.49 (s)	
$a = [mean, STD]$ $N = 10\%$	0.63 (s)	0.56 (s)	1.02 (s)	0.75 (s)	27.55 (s)	

Case Study II

The modified IEEE standard 69 bus distribution network (shown in Figure 4.10) is chosen for the second case study. The effectiveness of the proposed method is evaluated using this system. The system is suitably adapted to include a mix of commercial and residential loads and DGs. A set of experimental data (available for a time period of one year), obtained from Open Energy Information (OpenEI) [217], is utilized with the simulation time step of 1-hour. Three PVs are placed [218] on buses 21, 62, and 64 with maximum generation of 929.7 kW, 1075.2 kW, and 992.5 kW, respectively.

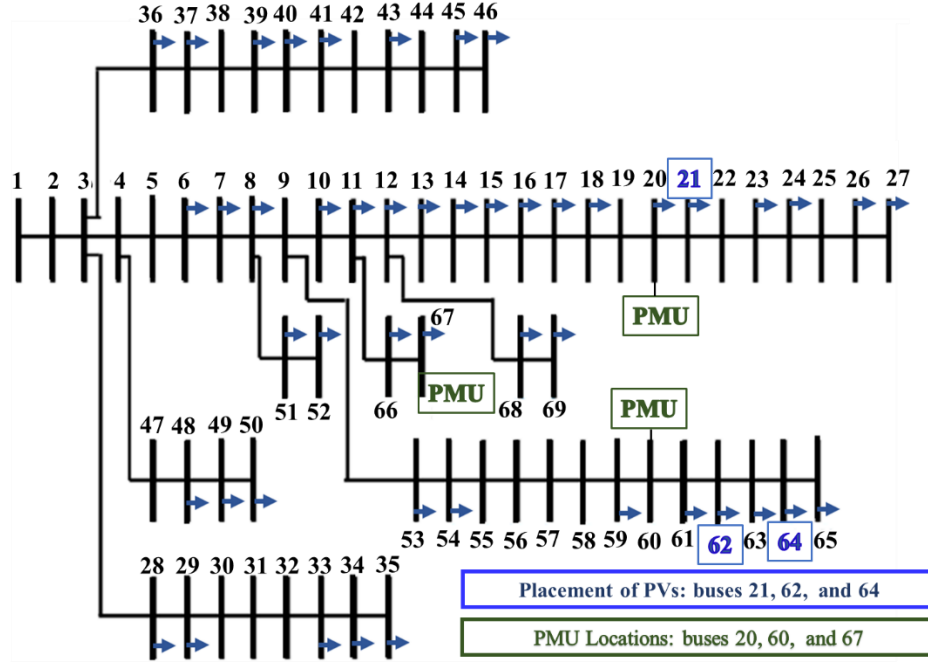


Figure 4- 10. The modified IEEE Standard 69 Bus Distribution System.

The hourly data of power generation of a photovoltaic system are computed and adopted based on actual data for Bozeman, MT, USA [213], [219].

- 1- Three PMUs—located at buses 20, 60, and 67 [245]—are assumed as the available measurement devices in the network. The bus voltage and branch current magnitudes from each PMU are used for FDIAs detection and performing DSSE using the proposed method. 8760 samples with a temporal resolution of 1 sample per hour for each PMU are available, resulting in data for one year.
- 2- Two types of attack vector are injected into available measurements:
 - a) $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$ [201]
 - b) $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$, where μ_z and σ_z are the average and standard deviation of each measurement vector, respectively. In this case, it is assume that the attacker knows the

distribution of each measurement and wants to falsify measurements based on the true measurement distribution.

- 3- All the data are then split into train and test sets with a ratio of 67% and 33%.
- 4- The pseudo-measurements of active and reactive power injections and flows are generated to make the system observable and perform WLS calculations with the inclusion of PMU measurements.
- 5- Voltage magnitudes and phase angles for all buses are considered as state variable: $x = [\delta_2, \dots, \delta_N, V_1, \dots, V_N]$, where δ_N , V_N are voltage phase and magnitude, respectively, and N is the number of buses. It is assumed that there are no measurement devices installed in the slack bus and $\delta_1 = 0$.
- 6- The standard deviation is considered as 50% of the nominal value for pseudo-measurements and 3% of the actual value of active and reactive power flow measurements. A gaussian error, with $3\sigma = 1\%$, is added to the PMU measurements (i.e., the voltage and branch current magnitudes) to model the uncertainty. To model zero injection buses, virtual measurements' error is considered as 10^{-8} .

False Data Injection Attacks on Measurements with $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$. In this case, the FDIAs are injected to measurement samples randomly, and 10% of each available measurement ($N_a = 876$) are falsified randomly by the FDIAs: $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

The results in Table 4.5 for the proposed and appropriate binary classification-only methods show that the proposed method successfully detected most FDIAs on each PMU measurement when the false data, $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$, are injected randomly to 10% of available measurements. The false data is the same for the independent binary classification method. The table shows that the accuracy and F1-score of the proposed method (0.954 and 0.5334, respectively) are better than when performing the binary classification only (0.916 and 0.521, respectively).

Table 4- 5. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification- only DNN model when $N_a = 876$, and $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

			Joint DNN Regression and Classification							DNN Classification Separately						
Actual Values	FDIA	$ V_{20} $	36	2	1	1	0	2	1	34	0	1	5	1	1	1
		$ V_{60} $	1	31	3	0	0	0	2	1	31	1	0	1	1	1
		$ V_{67} $	2	0	27	0	0	0	1	4	0	25	0	0	1	0
		$ I_{20-21} $	0	0	0	1	0	0	31	4	0	0	3	0	0	29
		$ I_{60-61} $	0	0	0	0	0	0	33	0	0	0	1	0	0	132
		$ I_{67-68} $	0	0	0	0	0	0	42	0	0	0	2	0	0	40
	No FDIA	0	0	0	3	3	3	2666	0	0	0	108	10	0	2557	
			$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $	No FDIA	$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $	No FDIA
		Predicted Values							Predicted Values							
		Accuracy = 0.9546 F_1 -Score = 0.5334							Accuracy= 0.9166 F_1 -Score = 0.5231							

The results of DSSE calculation from the proposed, the regression-only, and the WLS methods are shown in Figure 4.11. It is clear from this figure that the proposed method has a better performance in MPE and MAE criteria compared to other methods in the presence of bad data.

When no bad data is present, both the DNN-based methods (regression-only and combined regression and classification) have similar performances for state estimation. The WLS estimator has the worst performance in both cases, and the performance is significantly worse in the presence of the bad data measurements.

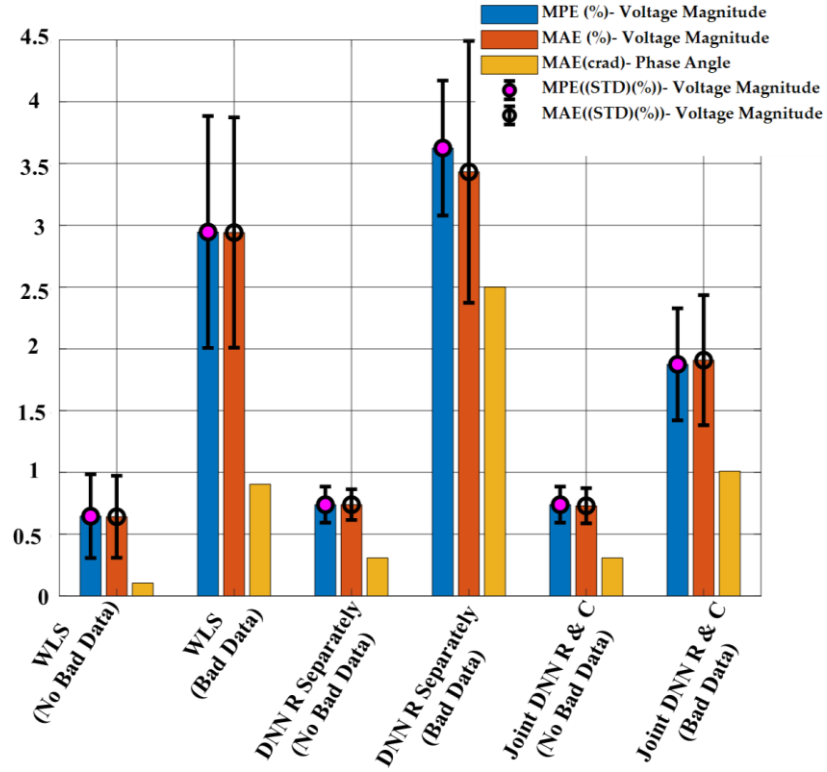


Figure 4- 11. DSSE results obtained using the proposed method, the Regression- only DNN Model, and WLS with or without Bad Data when $N_a = 876$, and $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

False Data Injection Attacks on Measurements with $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$. The results in Table 4.6 for the proposed and appropriate binary classification-only methods show the proposed method successfully detected most of FDIAs on each PMU measurement when false data $z_a \sim \mathcal{N}(\mu_z, \sigma_z)$ are injected randomly to 5% of available measurements, and it works better than an independent binary classification method.

The accuracy and F1-score of the proposed method (0.9723 and 0.722, respectively) are better than when performing binary classification only (0.970 and 0.702, respectively).

Table 4- 6. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification-only DNN model when $N_a = 438$ (5%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

			Joint DNN Regression and Classification						DNN Classification Separately							
Actual Values	FDIA	$ V_{20} $	22	2	2	2	0	0	4	22	3	1	2	0	0	4
		$ V_{60} $	0	18	0	1	6	1	3	0	17	0	2	4	1	5
		$ V_{67} $	0	0	25	1	0	4	7	1	0	24	1	0	2	9
		$ I_{20-21} $	0	0	1	21	0	0	9	0	2	1	18	0	0	10
		$ I_{60-61} $	0	2	0	0	18	1	10	0	2	0	0	17	2	10
		$ I_{67-68} $	0	0	0	0	1	8	22	0	1	0	0	0	9	21
	No FDIA	0	0	0	0	0	1	2699	0	0	0	0	0	1	2699	
			$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $	No FDIA	$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $	No FDIA
Predicted Values									Predicted Values							
Accuracy = 0.9723 F_1 -Score = 0.7224									Accuracy= 0.9705 F_1 -Score = 0.7021							

Figure 4.12 shows the results of the DSSE calculation from the proposed, the regression-only, and WLS methods. As is clear from the Figure, the proposed method, has a better performance in the MPE and MAE criteria compared to the other methods in the presence of bad data. This is similar to what was seen in the previous case study, when no bad data is present. Both DNN-based methods (regression-only and combined regression and classification) have similar performance for state estimation. The WLS estimator has the worst performance in both cases, and the performance is significantly worse in the presence of bad data measurements.

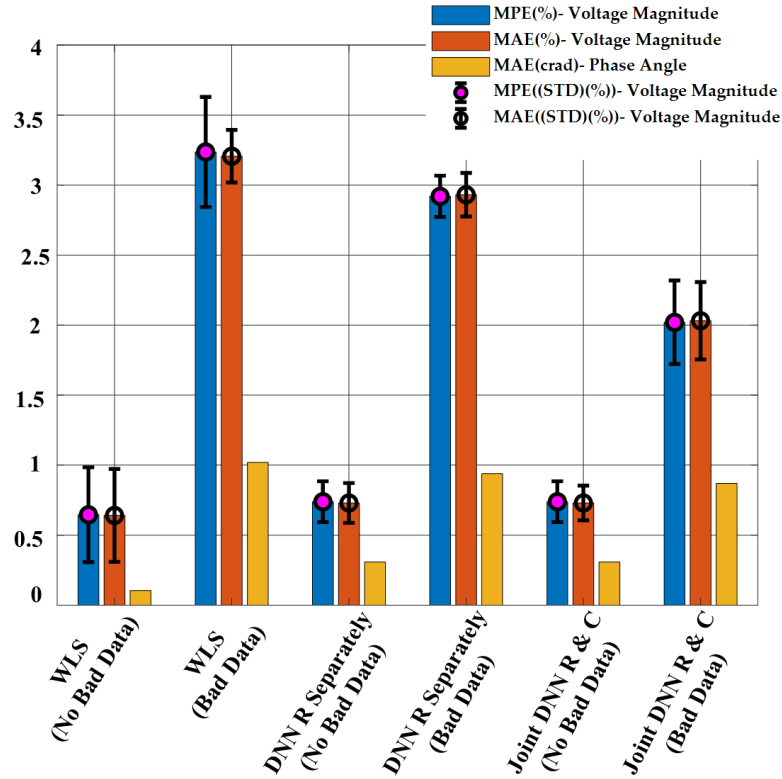


Figure 4- 12. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 438$ (5%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Table 4.7 shows the results for the proposed and appropriate binary classification-only methods, where the proposed method successfully detected most FDIAs on each PMU measurement when false data $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$ are injected randomly to 10% of available measurements, and it works better than an independent binary classification method. The accuracy and F1-score of the proposed method (0.9619 and 0.792, respectively) are also better than when performing the binary classification only (0.957 and 0.770, respectively).

Table 4- 7. Confusion matrix showing bad data detection results and accuracy values obtained using the proposed method and the binary classification- only DNN model when $N_a = 876$ (10%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Actual Values		Joint DNN Regression and Classification							DNN Classification Separately						
			$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $		$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $
FDIA	$ V_{20} $	64	1	1	2	1	0	4	64	0	2	1	1	0	5
	$ V_{60} $	0	47	1	0	9	0	4	0	45	2	0	9	1	4
	$ V_{67} $	0	3	53	4	3	0	1	0	5	52	2	4	0	1
	$ I_{20-21} $	1	1	1	52	0	0	4	1	2	0	52	0	0	4
	$ I_{60-61} $	0	10	0	0	47	1	13	0	14	0	0	41	3	13
	$ I_{67-68} $	0	0	0	0	0	13	41	0	0	0	0	0	13	41
No FDIA		0	0	0	0	0	4	2505	0	0	0	0	0	8	2501
		$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $	No FDIA	$ V_{20} $	$ V_{60} $	$ V_{67} $	$ I_{20-21} $	$ I_{60-61} $	$ I_{67-68} $	No FDIA
		Predicted Values							Predicted Values						
		Accuracy = 0.9619							Accuracy= 0.9574						
		F_1 -Score = 0.7921							F_1 -Score = 0.7709						

The results of DSSE calculation from the proposed, the regression-only, and WLS methods are shown in Figure 4.13. It is clear from this figure that the proposed method has a better performance in MPE and MAE criteria compared to the other methods in the presence of bad data. Once again, when no bad data is present, both DNN-based methods (regression-only and combined regression and classification) have similar performance for state estimation. The WLS estimator has the worst performance in both cases, and the performance is significantly worse in the presence of bad data measurements.

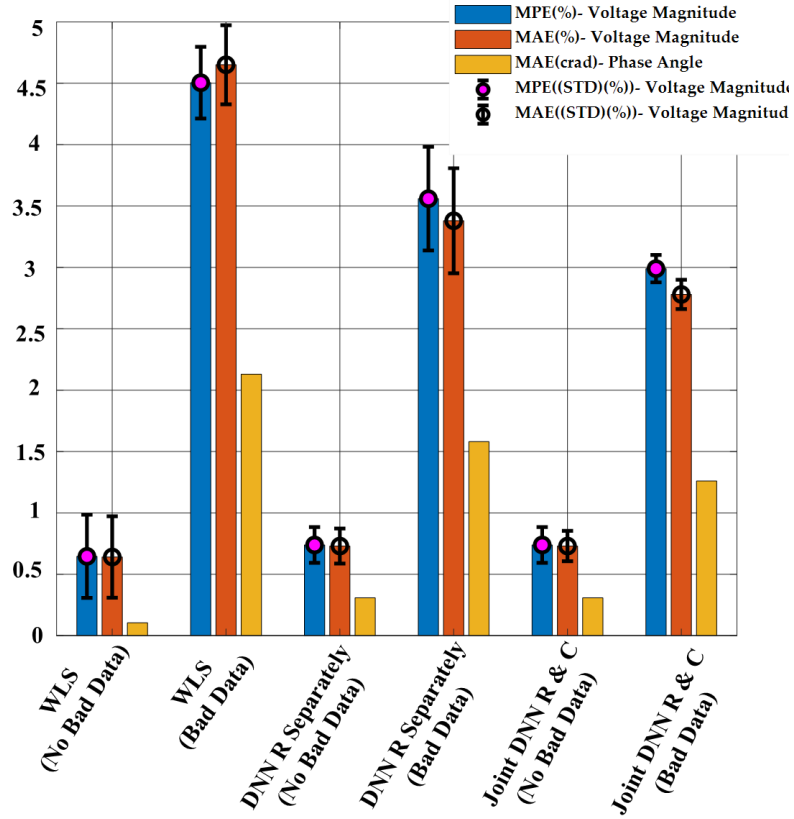


Figure 4- 13. DSSE results obtained using the proposed method, the regression- only DNN model, and WLS with or without bad data when $N_a = 876$, and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

In Table 4.8, the execution time for the proposed, regression- only, and WLS methods are shown when the testing data are the same for each method. It is clear from the table that the

execution time is decreased significantly when regression and binary classification are performed simultaneously by a joint DNN model. The WLS execution time is ten times that of the other two methods. As mentioned previously, each iteration recalculates the Jacobian matrix, which is based on physical parameters of a network, which increases the execution time. Therefore, by applying the proposed method, which is based on a data-based approach, the execution time is decreased significantly.

Table 4- 8. Execution time for the proposed, regression-only, and WLS methods

	Joint DNN Regression and Classification		DNN Regression Separately		WLS	
FDIA	$ V $	θ	$ V $	θ	$ V $	θ
$a = [\pm 90\%, 105\%]$	0.41 (s)	0.63 (s)	0.24 (s)	0.15(s)	43 (s)	
$a = [mean, STD]$ $N = 5\%$	0.24 (s)	0.23 (s)	0.15 (s)	0.15 (s)	51 (s)	
$a = [mean, STD]$ $N = 10\%$	0.17 (s)	0.15 (s)	0.18(s)	0.23 (s)	59 (s)	

Conclusion

A new method using a DNN approach is proposed to simultaneously perform DSSE calculation and FDIA detection on measurements in DNs. Voltage magnitudes and phase angles are defined as state vector variables. The proposed method considers the constraints of DG penetration and limitations on the installation of measurement tools in distribution networks, making it more applicable in a real-world setting. A single DNN model with two hidden layers is designed to perform both the regression (DSSE calculation) and binary classification (FDIA detection), and the results are compared when regression and binary classification are done with two separate DNN models using the same hyperparameters as the proposed DNN model.

Moreover, DSSE calculation based on the WLS method, along with PMU and pseudo-measurements, is performed to make a comparison between data-based and model-based approaches. In this work, it is shown that DSSE calculation can be performed precisely from corrupted measurements and simultaneously identify FDIAs also on corrupted measurements with high accuracy.

CHAPTER FIVE

DISTRIBUTION SYSTEM STATE ESTIMATION USING HYBRID TRADITIONAL AND
ADVANCED MEASUREMENT FOR GRID MODIRNIZATION

Due to the scarcity of measurement devices in distribution Networks, Distribution System State Estimation calculation is challenging [246]. In DNs, the lack of measurements causes issues with observability and the ability to perform DSSE computations. Generating pseudo-measurements, which are acquired from load forecasting based on Advanced Metering Infrastructure (AMI) measurements, is a common way to address these issues [13], [27]. Conventional measurements from Control and Data Acquisition (SCADA) measurements along with pseudo-measurements are used to perform SE calculation using low sample rates of SCADA measurements which are asynchronous and available with a long interval time [30]. Also, the conventional DSSE results are not precise enough since SCADA measurements and have low accuracy [31], [32].

New features—including DGs, two-way communications, electric vehicles, variable loads, etc.—will increase the need for accurate DSSE calculations [49], [143]. Therefore, installing smart measurement devices, such as PMUs and micro PMUs (μ PMUs), are currently involved in grid modernization projects to improve DSSE results [33], [34]. PMUs are extensively used in Transmission Networks (TNs) [35]. However, because DNs have different characteristics than TNs, are PMUs are not as extensively used in DNs [36].

Traditional and Smart Measurements

The penetration of DGs, EVs, and variable loads bring the complex interaction characteristics with high uncertainties that makes the operation, monitoring, and management of DNs more challenging [156], [247]. SCADA measurement systems are currently applied to monitor and control the DNs [248]. The information from Remote Terminal Units (RTUs), which are conventional measurement devices, are installed at various locations in DNs, are recorded by SCADA [249]. Measurement data are unsynchronized voltage and branch current magnitudes, real and reactive power injections, and flows within a lowest refresh interval time [250]. These measurements, available using SCADA systems, fail to capture the dynamic behavior of ADNs [251]; therefore, PMUs are introduced to facilitate a suitable and reliable monitoring of ADNs in terms of higher precision and speed.

The Phasor Measurement Units (PMUs) are advanced measurement devices in power grids to provide an accurate and synchronized voltage bus and branch current phasor measurement [252]. PMUs have been widely used for wide-area monitoring, protection, and control of TNs [26]. However, PMUs have not been deployed widely in ADNs due to economic constraints as well as structure of the DNs [253], [254]. DN monitoring will be enhanced by incorporating more PMUs in an ongoing modernization effort for ADNs due to the following reasons [255], [256]:

- The linear mathematical SE calculation will be formulated since state variables can be measured by PMUs.

- PMU measurements are available with a high precision (0.1% and 0.01 rad for magnitudes and phase angle, respectively), thus SE computations can be performed using PMU measurements with high-speed processing and accuracy.
- Tracking the dynamic behaviors of ADNs could be captured using high resolution PMU data in DSSE calculations.

SE results will be enhanced by Including different measurements data from multiple measuring instruments due to the redundancy [254]. But measuring information integration from different sources is not straightforward because of two main reasons [257]:

- 1- The absence of unique coordination time among different measurements leads to time-skew problem, and datasets might not be established in an exact time representative in a control center. Time inconsistency might also be resulted owing to the effect of communication delays of monitoring devices [258].
- 2- The precision weights of measured quantities from SCADA and PMU measurements are different. Therefore, current SE software should be altered or corrected to make all measurement values applicable in SE algorithms [259].

The SCADA, and PMU measurements, and their refresh samples are shown in Figure 5.1(a), and (b), respectively.

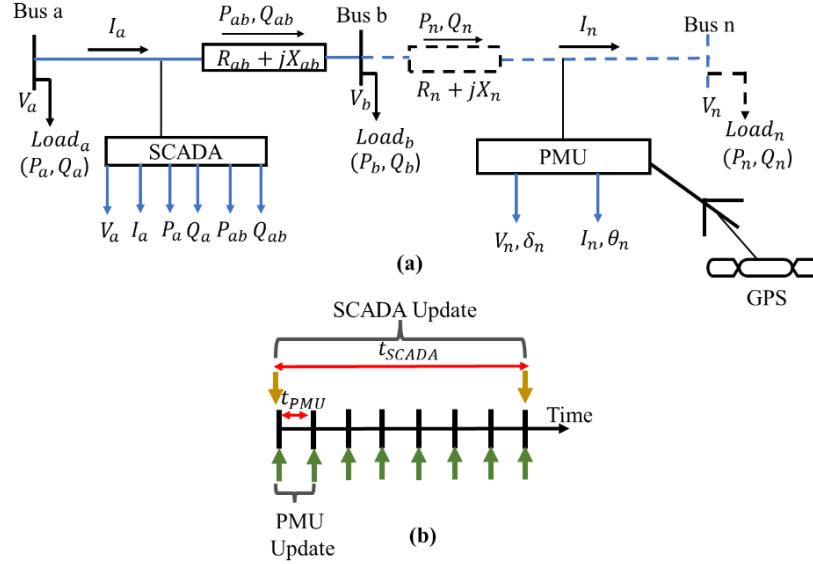


Figure 5- 1. (a) SCADA and PMU measurements in a power network, (b) SCADA and PMU measurements refresh rates.

Distribution System State Estimation

SE calculation is the function of Energy Management System (EMS) in a power grid. SE calculations make a way to monitor real-time power grid operation and develop more power grid analysis including power flow calculations, power dispatching, observability analysis, bad-data detection, reliability assessment, etc. DSSE strategies should be designed according to needs and requirements of DNs.

The first DSSE technique was established when the SCADA system was introduced in 1990 [260]. Traditional SE techniques are commonly performed based on conventional measurements which are commonly obtained from SCADA systems [58]. However, since SCADA measurements are not enough to acquire for SE calculations, pseudo-measurements are considered along with traditional measurements to make a system observable [27].

In order to calculate the state estimation variables, the WLS method is generally applied on both TNs and DN [260]. Using WLS, the state vector can be defined based on node voltage [261], [262] or branch current variables [263], [264]. When state variables are calculated, other network parameters can also be determined.

In the node voltage definition, the bus voltage phase angle as well as voltage magnitude are considered as a state vector [265]: $x = [\theta, V]$. When only conventional measurements are available, and no measurement devices are placed at the slack bus the state vector is defined in polar coordination as follows:

$$x = [\theta_2, \dots, \theta_N, V_1, \dots, V_N] \quad (1)$$

Where, N is the number of buses. The slack bus is considered as a reference, $\theta_1 = 0$.

When a PMU is located at the slack bus, the phase angle at the slack bus is not zero, and it should be involved in the state vector, and the state vector is represented as:

$$x = [\theta_1, \dots, \theta_N, V_1, \dots, V_N] \quad (2)$$

The relationship between the state variables and the measurement data is called the measurement function, represented as:

$$z = h(x) + e \quad (3)$$

where $z = [z_1, \dots, z_M]^T$ is the measurement vector, and M is the number of measurements. Bus voltage magnitudes and phase angles, branch current magnitudes and phases as well as active and reactive power flows and injections can be involved in the measurement vector.

The equations for standard active (P_i) and reactive (Q_i) power injections at bus i are defined:

$$P_i = V_i \sum_{j=0}^N V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}), \quad (4)$$

$$Q_i = V_i \sum_{j=0}^N V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}), \quad (5)$$

G_{ij} and B_{ij} are the real and imaginary part in a nodal admittance matrix element Y_{ij} and $\theta_{ij} = \theta_i - \theta_j$ is the standing phase angle difference between buses i and j .

The active (P_{ij}) and reactive (Q_{ij}) power flow from bus i to bus j are expressed as:

$$P_{ij} = V_i V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) - G_{ij} V_i^2 \quad (6)$$

$$Q_{ij} = V_i V_j (G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij}) + B_{ij} V_i^2 \quad (7)$$

The relationship between the measurement vector and the state variable $x = [x_1, \dots, x_N]^T$ (N is the number of state variables) is shown by $h = [h_1, \dots, h_M]^T$. h is the list of measurement functions which is commonly nonlinear. $e = [e_1, \dots, e_M]^T$ is the vector measurement error which follows the Gaussian distribution $e \sim \mathcal{N}(0, W)$, where W is the covariance matrix of the measurement errors $W = \text{diag}\{\sigma_{z_1}^2, \sigma_{z_2}^2, \dots, \sigma_{z_m}^2\}$, and zero mean.

The WLS technique is commonly applied to solve the SE problem. In this method, state vectors are obtained by minimizing the sum of the residual squares as follows.

$$\hat{x} = \arg \min_x J(x) = \arg \min_x ((z - h(x))^T W^{-1} (z - h(x))). \quad (8)$$

The Iterative Gauss-Newton numerical method is applied to solve the optimization problem (8) with the objective function $J(x)$.

$$[G(x^k)] \Delta x^k = H^T(x^k) W^{-1} [z - h(x^k)],$$

$$x^{k+1} = x^k + \Delta x^k, \quad (9)$$

$$G(x) = \frac{\partial J(x)}{\partial x} = H^T(x) W^{-1} H(x),$$

Where G is the gain matrix, and Δx^k is the updated state vector to determine a new vector x^k .

The iterative calculation proceeds until the preset convergence criterion is attained. The prefix

threshold ε is compared to the largest absolute value of the updated state vector (Δx^k), and the state variables will be determined in the last iteration when $\max(|\Delta x|) < \varepsilon$.

$H(x) = \partial h(x)/\partial(x)$ is the Jacobian matrix which is derived from the function vector $h(x)$.

$$H(x) = \frac{\partial h(x)}{\partial(x)} = \begin{bmatrix} \frac{\partial h_1(x)}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \dots & \frac{\partial h_1(x)}{\partial x_n} \\ \frac{\partial h_2(x)}{\partial x_1} & \frac{\partial h_2(x)}{\partial x_2} & \dots & \frac{\partial h_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_m(x)}{\partial x_1} & \frac{\partial h_m(x)}{\partial x_2} & \dots & \frac{\partial h_m(x)}{\partial x_n} \end{bmatrix} \quad (10)$$

The measurement function definitions and the Jacobian matrix in the SE computations are different when either traditional or smart measurements are included in the measurement vector [266]. In (11), the Jacobian Matrix is shown when the traditional ($H_{Traditional}$) and smart (H_{Smart}) measurements are included in the measurement vector, respectively.

$$H_{Traditional} = \begin{bmatrix} \frac{\partial V_{SCADA}}{\partial \theta} & \frac{\partial V_{SCADA}}{\partial V} \\ \frac{\partial P_{inj}}{\partial \theta} & \frac{\partial P_{inj}}{\partial V} \\ \frac{\partial Q_{inj}}{\partial \theta} & \frac{\partial Q_{inj}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \theta} & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{flow}}{\partial \theta} & \frac{\partial Q_{flow}}{\partial V} \end{bmatrix} \quad H_{Smart} = \begin{bmatrix} \frac{\partial V_{SCADA}}{\partial \theta} & \frac{\partial V_{SCADA}}{\partial V} \\ \frac{\partial V_{PMU}}{\partial \theta} & \frac{\partial V_{PMU}}{\partial V} \\ \frac{\partial I_{PMU}}{\partial \theta} & \frac{\partial I_{PMU}}{\partial V} \\ \frac{\partial I_{imag}}{\partial \theta} & \frac{\partial I_{imag}}{\partial V} \\ \frac{\partial P_{inj}}{\partial \theta} & \frac{\partial P_{inj}}{\partial V} \\ \frac{\partial Q_{inj}}{\partial \theta} & \frac{\partial Q_{inj}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \theta} & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{flow}}{\partial \theta} & \frac{\partial Q_{flow}}{\partial V} \end{bmatrix} \quad (11)$$

Methodology

As mentioned earlier, SCADA measurements are commonly available from the beginning operation of ADNs to perform DSSE calculations and make them observable [267]. The

installation of PMU technology has gradually been developed as parts of grid modernization projects to move from traditional DNs to smart DNs [268]. Thus, DSSE results will be improved by incorporating PMU measurements as well as SCADA measurements.

Unfortunately, SCADA measurements typically have a slower refresh rate compared with the newly installed PMU measurements [269]. With smart DNs, it is important to perform DSSE calculations as frequently as possible in order to understand the dynamic state of the system [270]. Rather than only performing SEs as each new SCADA measurement arrives, we want to perform DSSE calculations with every new PMU measurement. This requires a method of interpolating SCADA measurements so they can be incorporated into the DSSE algorithm. Therefore, it is assumed that an ADN is equipped with SCADA measurements. After a period of time, the ADN is updated to include PMUs as part of a grid modernization procedure.

Figures 5.2 and 5.3 illustrate how a regression model is trained to predict future SCADA values prior to and after installing PMUs. Regression models are a category of supervised machine learning, where the corresponding output (y) at each input (x) is available [271]. In the training phase, a relationship between x and y is determined and expressed as f ; i.e., $y = f(x)$. A regression model is chosen in this chapter, since there are enough data points to train a model.

Note that the regression model is only used for DSSE after PMUs are installed in order to produce interpolated SCADA measurements that can be used to perform DSSE more frequently. In this chapter, the SCADA prediction model is pre-trained prior to PMU installation and regularly updated the model after the PMU installation.

Prior to PMU installation, DSSE calculations based on WLS method are performed using SCADA and pseudo-measurements. A multi-output regression model is trained based on these

DSSE results to predict an interpolated SCADA measurement, as shown in Figure 5.2. The ground truth used for the interpolated SCADA measurement is obtained using backward-forward power flow calculation. This procedure is repeated for every SCADA measurement that arrives prior to the PMU installation.

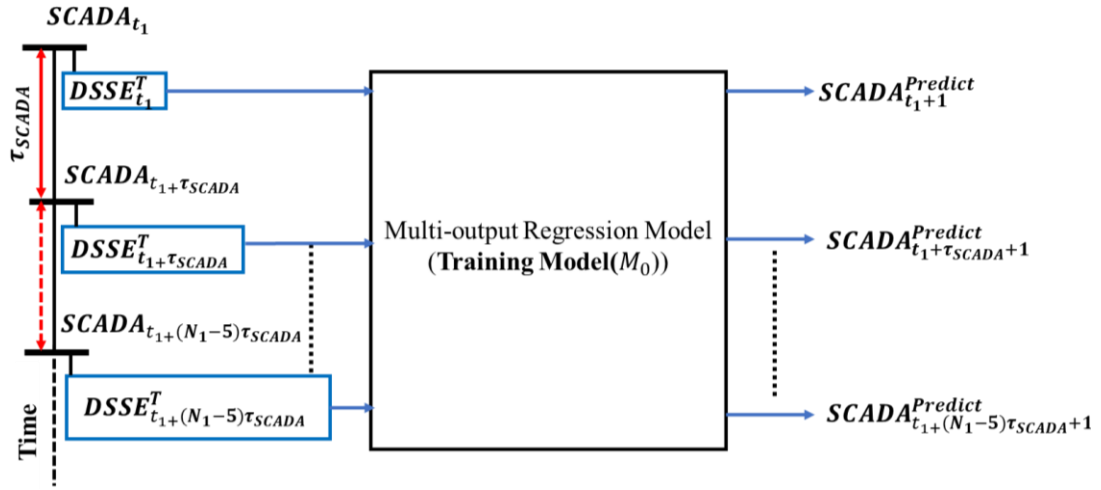


Figure 5- 2. A Multi-output regression model to predict SCADA measurements. For the figure, it is assumed that the PMUs are installed in the DN at time step N_1 , so the procedure is performed through time step N_1-5 .

A different training procedure will be performed after adding the PMUs to the ADN. In the updated procedure, SCADA measurements are predicted using the most recent DSSE calculations, which in turn depend on the most recent PMU measurement and the most recent SCADA measurement or prediction. As shown in Figure 5.3, a training model will be updated every time real SCADA measurements are provided in order to improve model performance and make it robust against dynamic changes of ADNs. For instance, when SCADA and PMU measurements are refreshed at t_1 , a training model is updated from M_0 to M_1 .

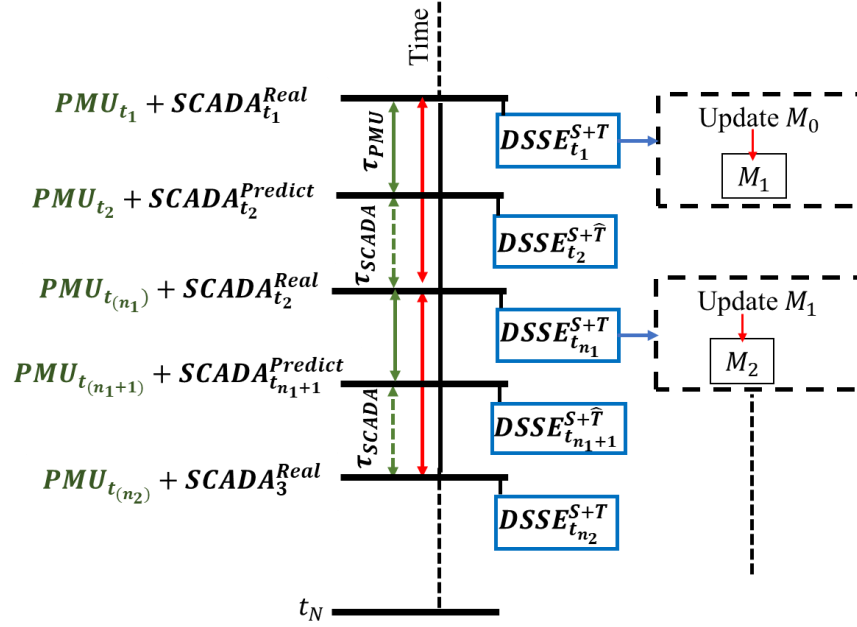


Figure 5- 3. DSSE calculations and updating a training model after PMU installation.

In order to evaluate the DSSE result, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are calculated using (12) and (13), respectively:

$$MAE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{x}_i - x_i}{x_i} \right| \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{x}_i - x_i)^2}{N}} \quad (13)$$

In the above equations, $\hat{x}$ is the estimated value, x is the actual value, and N refers to the data set size.

The interface between MATLAB (2022) and Anaconda-Python is used to train and update the model using the DSSE results obtained from the SCADA and PMU measurements. All tests were performed on a desktop computer with an Intel Core i5 processor clocked at 1.60 GHz with 8.00 GB RAM.

Results

Case Study I

The effectiveness of the proposed method is evaluated on the modified IEEE 33-bus distribution network as one of the case studies (shown in Figure 5.4). Actual values of the DN are calculated based on power flow calculations.

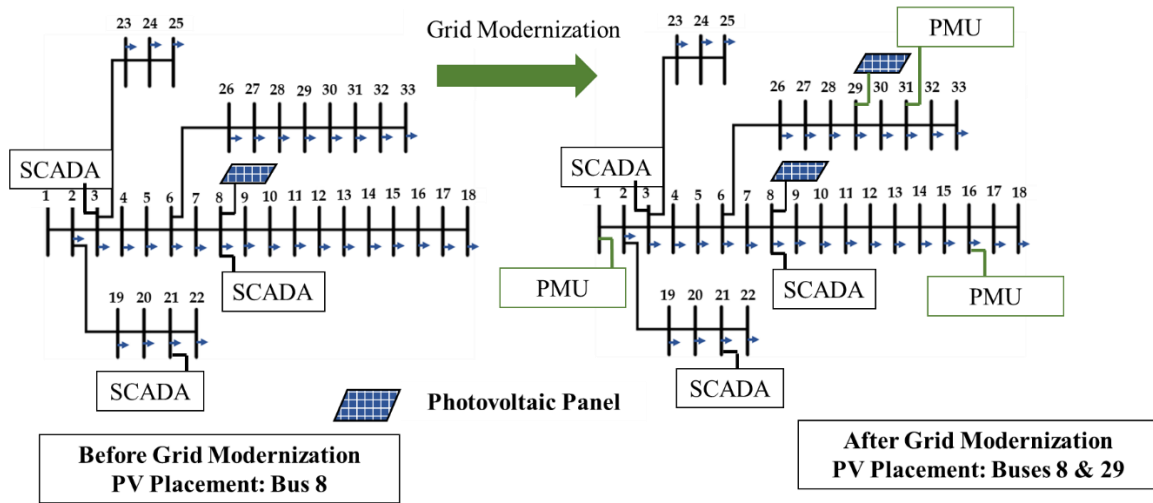


Figure 5- 4. IEEE standard 33 bus distribution system before and after grid modernization. Prior to modernization, there is a PV on bus 8 and SCADA units on buses 3, 8, and 21. After modernization, an additional PV is installed on bus 29 and PMU devices are installed on buses 1, 16, and 31.

The following assumptions are made about the system prior to PMU installation:

- 1- Measurements are randomly calculated based on their probability density function for each Monte Carlo trial.
- 2- A Gaussian distribution, with $3\sigma = 50\%$ of the nominal value, is considered for power injection on the buses [272].

- 3- The PV panel placed at bus 8 [273] with the maximum generation capacity of 104 kW (The PV is modeled as a negative load), and three SCADA devices are located at buses 3, 8, and 21. Voltage magnitudes, active and reactive power flows, and injections are considered as real measurements ($|V|_3, |V|_8, |V|_{21}$),
 $(P_{3-4}^{flow}, P_{3-23}^{flow}, P_{8-9}^{flow}, P_{21-22}^{flow}), (Q_{3-4}^{flow}, Q_{3-23}^{flow}, Q_{8-9}^{flow}, Q_{21-22}^{flow}), (P_3^{inj}, P_8^{inj}, P_{21}^{inj})$, and
 $(Q_3^{inj}, Q_8^{inj}, Q_{21}^{inj})$ from SCADA to perform DSSE calculations.
- 4- The pseudo-measurements of active and reactive power injections and flows are generated to make the system observable and to perform the WLS calculations.
- 5- The standard deviation is considered as 50% of the nominal value for pseudo-measurements, 2% of the voltage magnitude and 3% of the actual value of active and reactive power flow and injection measurements [149].
- 6- Voltage magnitudes and phase angles for all the buses are considered as a state variable: $x = [\delta_2, \dots, \delta_N, V_1, \dots, V_N]$, where δ_N, V_N are the voltage phase angle and magnitude, respectively, and N is the number of buses. It is assumed that there are no measurement devices installed in the slack bus and $\delta_1 = 0$. The Jacobian matrix will be formed as $H_{Traditional}$, which is shown in section 5.2.
- 7- The SCADA measurements are refreshed every 5 iterations, and the DSSE calculations are performed using the SCADA and pseudo-measurements every 5 iterations, concluding at $N_{MC} = 5995$. It is assumed that after this Monte Carlo trial, PMUs are added to the system.
- 8- The DSSE results will be used for training a model to interpolate the SCADA measurements, as discussed in section 5.3. It should be noted that the inputs are corrupted

by random Gaussian error, with $3\sigma = 2\%$ and $3\sigma = 1 \text{ crad}$ for voltage magnitude and phase angle values, respectively, before feeding to a training model. This initial model will be used to predict the SCADA measurements after the installation of PMUs, to enable the DSSE calculations to be performed more frequently.

- 9- The WLS calculations were performed using MATLAB and the iteration process is stopped when the minimum difference between two iterations is $\varepsilon = 10^{-6}$.

The following assumptions are made about the system after the PMU installation:

- 1- Three PMUs are placed at buses 1, 16, and 31 [273], and one more PV with maximum generation capacity of 80kW is added to a DN at bus 29 at $N_{MC} = 6000$ to meet the new demands of the DN in grid modernization development.
- 2- To model the uncertainty of the PMU measurements, a Gaussian error with $3\sigma = 1\%$ and $3\sigma = 1 \text{ crad}$ is added to the voltage and branch current magnitudes, and voltage and branch current phase angles, respectively [274].
- 3- Since a PMU is installed in the slack bus, the phase angle at bus 1 is included in the state vector, so the full state vector is defined as: $x = [\delta_1, \dots, \delta_N, V_1, \dots, V_N]$.
- 4- It is assumed that the PMU measurements are updated at each iteration, i.e., $\tau_{SCADA} = 5\tau_{PMU}$, and at each iteration SCADA measurements are predicted from a training model.
- 5- The training model is updated every time N_{MC} is a multiple of 20 (6020, 6040, 6060, etc.) using the DSSE results obtained from the SCADA and PMU measurements during that time.

- 6- In order to verify the performance of the proposed method, the DSSE results are compared to when the SCADA measurements are replaced using a sample-and-hold technique.
- 7- The Estimated value from WLS calculations and the actual value from the power flow calculations are compared in MAE and RMSE criteria, as shown in Table 5.1.

Table 5- 1 DSSE results of voltage magnitudes and phase angles after grid modernization.

Criteria	Voltage Magnitude (<i>p. u.</i>)		Phase Angle (<i>rad</i>)	
	Proposed Method	Sample-and-hold technique	Proposed Method	Sample-and-hold technique
MAE	8.2×10^{-4}	9.7×10^{-4}	0.0060	0.0071
RMSE	1.02×10^{-5}	1.45×10^{-5}	7.9×10^{-5}	9.2×10^{-5}

It is clear from Table 5.1 that the MAE and RMSE values for the state variables (voltage magnitude and phase angle) are improved from the proposed method compared to the sample-and-hold technique. This means that predicting the SCADA measurements based on updating the training model has a better performance. This is true even after more PVs are introduced to the DN, showing that the proposed method is robust in a more dynamic environment.

Case Study II

The effectiveness of the proposed method is also evaluated using the modified IEEE standard 69-buse distribution network (shown in Figure 5.5). The system is suitably adopted to include a mix of commercial and residential loads and DGs. A set of experimental data (available for a time period of one year), obtained from Open Energy Information (OpenEI) [217], is

utilized with the simulation time step of 1h. The hourly data of power generation of a PV system are computed and adopted based on the actual data for Bozeman, MT, USA [213], [219].

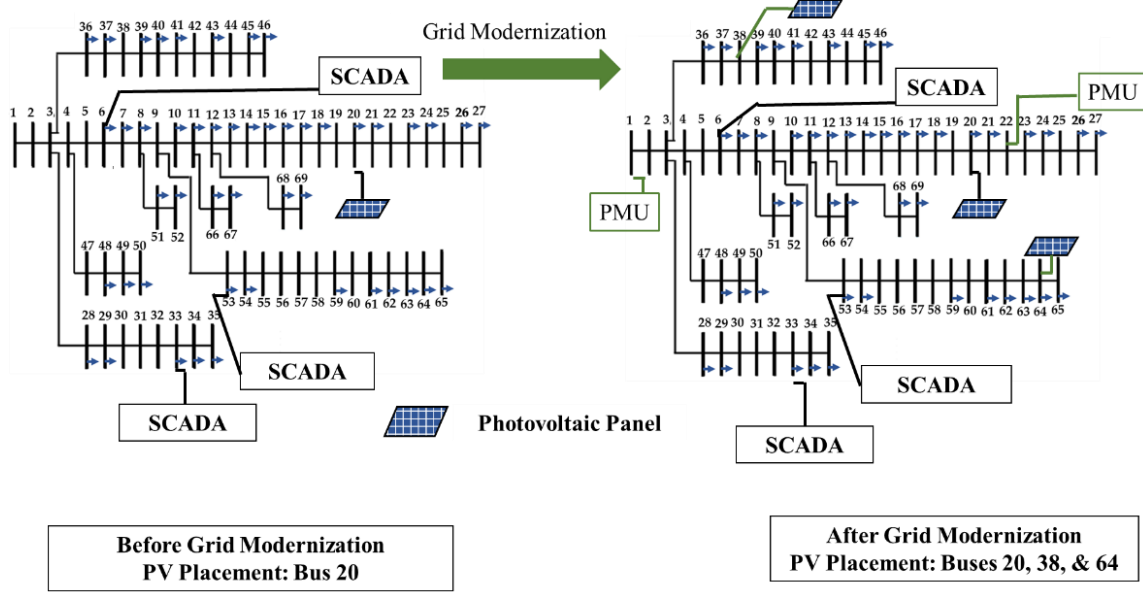


Figure 5- IEEE standard 69 bus distribution system before and after grid modernization. Prior to modernization, there is a PV on bus 20 and SCADA units on buses 33, and 53. After modernization, two additional PVs are installed on bus 38 and 64, and PMU devices are installed on buses 1, and 22.

The following assumptions are made about the system prior to PMU installation:

- 1- It is assumed that one PV is placed at bus 20 [273] with a maximum generation capacity of 15.33 kW (The PV is modeled as a negative load), and three SCADA devices are located at buses 6, 33, and 53. Voltage magnitudes, active and reactive power flows, and injections are considered as real measurements

$(|V|_6, |V|_{33}, |V|_{53}), (P_{6-7}^{flow}, P_{33-34}^{flow}, P_{53-54}^{flow}), (Q_{6-7}^{flow}, Q_{33-34}^{flow}, Q_{53-54}^{flow}), (P_6^{inj}, P_{33}^{inj}, P_{53}^{inj}),$
and $(Q_6^{inj}, Q_{33}^{inj}, Q_{53}^{inj})$ from SCADA to perform the DSSE calculations.

- 2- The pseudo-measurements of active and reactive power injections and flows are generated to make the system observable and to perform WLS calculations.
- 3- The standard deviation is considered as 50% of the nominal value for pseudo-measurements, 2% of the voltage magnitude and 3% of the actual value of active and reactive power flow and injection measurements [149].
- 10- Voltage magnitudes and phase angles for all the buses are considered as a state variable: $x = [\delta_2, \dots, \delta_N, V_1, \dots, V_N]$, where δ_N, V_N are the voltage phase angle and magnitude, respectively, and N is the number of buses. It is assumed that there are no measurement devices installed in the slack bus and $\delta_1 = 0$. The Jacobian matrix will be formed as $H_{Traditional}$ which is shown in section 5.2.
- 4- SCADA measurements are refreshed every 5 iterations, so the DSSE calculations are performed using SCADA and pseudo-measurements every 5 iterations, concluding at $t = 4125$. (After this time, PMU devices are added to the system.)
- 5- The DSSE results will be used for training a model to interpolate SCADA measurements as discussed in section 5.3. This initial model will be used to predict SCADA measurements after the installation of PMUs to enable the DSSE calculations to be performed more frequently.
- 6- The WLS calculations were performed using MATLAB and the iteration process is stopped when the minimum difference between two iterations is $\varepsilon = 10^{-6}$.

The following assumptions are made about the system after PMU installation:

- 1- Two PMUs are placed at buses 1 and 22 [273] and two more PVs are added to a DN at buses 38 and 64 with maximum generation 20.83kW and 16.67 kW, respectively at $t = 4130$ to meet the new demands of DNs in grid modernization development.
- 2- To model the uncertainty of the PMU measurements, a Gaussian error, with $3\sigma = 1\%$ and $3\sigma = 1 \text{ crad}$ is added to voltage and branch current magnitudes, and voltage and branch current phase angles, respectively [274].
- 3- Since a PMU is installed in the slack bus, phase angle at bus 1 is included in the state vector and is defined as: $x = [\delta_1, \dots, \delta_N, V_1, \dots, V_N]$.
- 4- It is assumed that the PMU measurements are updated at each iteration, i.e., $\tau_{SCADA} = 5\tau_{PMU}$, and at each iteration SCADA measurements are predicted from a training model.
- 5- The training model is updated once or twice per day using the DSSE results obtained from the SCADA measurements and PMU measurements during that time. The PMU measurement resolution is 20 measurements per day, so the model is updated every 10 (twice daily updates) or 20 (once daily updates) PMU measurements.
- 6- In order to verify the performance of the proposed method performance, the DSSE results are compared when SCADA measurements are replaced using a sample-and-hold technique.
- 7- The Estimated value from the WLS calculations and the actual value from the power flow calculations are compared in MAE and RMSE criteria and shown in Table 5.2.

Table 5- 2. DSSE Results of Voltage Magnitudes and Phase Angles After Grid Modernization when the training model is updated every 10- or 20-time steps (twice or once per day).

Criteria	Voltage Magnitude (<i>p. u.</i>)			Phase Angle (<i>rad</i>)		
	Proposed Method		Sample-and-hold technique	Proposed Method		Sample-and-hold technique
	Daily Update	Twice Daily Update		Daily Update	Twice Daily Update	
MAE	0.0031	0.0023	0.0035	0.0069	0.0062	0.0071
RMSE	4.8×10^{-5}	3.7×10^{-5}	5.2×10^{-5}	9.1×10^{-5}	8.5×10^{-5}	9.8×10^{-5}

This case study is based on the actual data, and it has more severe dynamic behaviors. When the model is updated daily, the results are improved, but still close to those obtained using the sample-and-hold technique. In order to improve the results, the training model is updated twice a day to become robust against the dynamic behavior of the DN. As it is clear from Table 5.2, the MAE and RMSE values for the state variables (voltage magnitude and phase angle) are improved using the proposed method when the model is updated twice a day compared to the sample-and-hold technique.

Conclusion

In this chapter, a new approach based on a regression model is proposed to utilize all the available measurements to improve the DSSE results. DSSE results have widely been dependent on traditional measurements before installation of smart measurement devices in ADNs. As part of grid modernization efforts toward an advanced grid, smart measurement devices have progressively been installed to manage and control the dynamic behaviors of ADNs. However,

all traditional measurement devices cannot be substituted with advanced ones over a short time, and it is essential to utilize different kinds of measurements using an appropriate method.

Conventional measurements have low update rates unlike smart measurement devices which collect measurements at high frame rates. The difference in frame rates requires interpolation between the traditional measurements to allow for synchronization between the traditional and advanced sensors.

The WLS method is initially performed using the traditional measurements to calculate the state variables; i.e bus voltage magnitudes and phase angles. Then, the network configuration is modified toward a modernized DN by installing smart measurement devices to improve the DSSE results and observability analysis. Since traditional and smart measurements have different sampling rates, the DSSE results using a traditional measurement before the smart measurement devices are added to the system, and their data are used to train a machine learning-based regression approach. This training model is used to predict the traditional measurements when these measurements are not available between their refreshing times after the installation of smart measurement devices.

The proposed regression method works effectively for the DSSE calculations, but it requires the model to be updated regularly in order to be robust in a dynamic environment. In this work, the results of the proposed method are compared with those obtained when the traditional measurements are incorporated into the DSSE calculation using a sample-and-hold technique after smart measurement installation are compared. The effectiveness of the proposed method is validated using the two case studies in the presence of DGs:

The DSSE results given illustrate that the proposed method performs better than the sample-and-hold method after advanced measurement installation. The proposed method could be used in advancing distribution grid modernization to enhance DN performance and capabilities.

CHAPTER SIX

CONCLUSION

Because of the deployment of new features and technologies, Distribution Networks (DNs) have been developed and upgraded to advanced networks. These modern technologies include renewable energy resources and storage devices which transform traditional infrastructures into smart ones. Even though the performance of a DN is improved and customer needs and wants are better acquired, new challenges and issues have emerged regarding DN control and management. For example, the degree of uncertainty in dynamic behaviors are intensified by the penetration of Distributed Generations (DGs). Likewise, the risk of cyber-attacks is also increased with the development of communication and advanced measuring devices in the smart DN.

To address these challenges, Distribution System State Estimation (DSSE) calculations have been introduced to obtain accurate real-time state information to enable reliable operation and control. DSSE methods are mostly developed based on specific DN features and characteristics since Transmission System State Estimation (TSSE) cannot be applied to DNs.

DSSE calculations are highly dependent on the number of real-time measurements from measurement instruments, but due to financial constraints and particular characteristics of DNs, installation of measurement devices is not possible at all buses in a DN. To solve this problem, generating pseudo-measurements from historical load and generation information has been established to address the lack of real measurements in a DN. Pseudo- and traditional measurements are fed in to DSSE calculations to obtain DN states. Unfortunately, DSSE results

from conventional and pseudo-measurements are not accurate enough to resolve DN control and decision problems.

Recently, as part of grid modernization procedures, limited numbers of smart measurement devices are being installed in DNs. This enhances DSSE performance and addresses the lack of accurate traditional and pseudo-measurements in Active Distribution Networks (ADNs). Smart and conventional measurement data are available with different sampling rates and precision and these measurements should be combined truly to acquire an accurate DSSE result and take advantage of all measurements in a DN.

Having adequate and accurate measurement data are two main conditions required to obtain precise DSSE results. However, measurement data integrity could be degraded by False Data Injection Attacks (FDIAs) on measurement information. If FDIAs cannot be identified and removed using an appropriate method, DSSE results could be influenced significantly, and a proper decision cannot be made by a control center to improve efficiency, reliability, resiliency, and security of a DN.

Machine learning approaches have already been extensively applied for different purposes in the DN, including power flow and DSSE calculations, cyber-security awareness, reliability and resilience assessments, and energy market analysis. Fast speed and high precision are two features of machine learning algorithms in contrast to model-based methods for DN applications. Machine learning methods are categorized into supervised, semi-supervised, unsupervised, and reinforcement learning.

In this study, different supervised machine learning methods are applied for generating pseudo-measurements, performing state estimation calculation and FDIAs detection simultaneously, and combining traditional and smart measurements toward grid modernization.

A limited number of available real measurements, as well as temporal and meteorological information, are used to generate pseudo-measurements considering dynamic behavior of photovoltaic panels during clear and cloudy sky conditions. Three Deep Neural Networks (DNNs) are considered as base learners, and the base learner results are applied to train a random forest to get the final results.

Pseudo-measurement results along with available real measurements are used to perform WLS calculations to verify the accuracy of the proposed method to generate pseudo-measurements. Moreover, the proposed method precision is compared to random split method in different criteria to verify the ability of the proposed method for generation pseudo-measurements.

A new method is proposed to perform DSEE calculation and FDIA detection simultaneously using a DNN method. In this method, regression (DSSE calculation) and classification (FDIAs detection) processes are performed at the same time at the output layer of the DNN. Two kinds of attack vectors are modeled and injected into Phasor Measurement Unit (PMU) data to evaluate the robustness, accuracy, and speed of the proposed method in performing state estimation calculations and detect FDIAs in DNs.

WLS method is also performed to assess the proposed method's robustness in state estimation calculation and FDIA detection. These results show that simultaneous state estimation

calculations and FDIA detection outperform the model-based performance in terms of accuracy and execution time in DNs.

A new method is proposed to hybridize traditional and advanced measurements using an online machine learning model. In this work, it is assumed that an ADN has been monitored using traditional measurements and the WLS method is performed to obtain DSSE results. The voltage magnitudes and phase angles at each bus are considered as the state vector. After a period of time, a network is modified by the installation of an advanced measurement device (such as a PMU) to facilitate ADN monitoring and control with a desired performance.

This work proposes a method for taking advantage of all available measurements to improve DSSE results. First, a machine learning-based regression model is trained from DSSE results obtained using only the traditional measurements available before the installation of smart measurement devices. After smart measurement devices are added to the network, the model predicts traditional measurements when these measurements are not available to enable synchronization between the traditional and smart sensors, despite their different refresh rates.

The results show that the regression model has improved performance under the condition that it continues to be updated regularly as more data are collected from the measurement devices. In this way, the SML model becomes robust and improves the DSSE performance even in the presence of more DGs. The results of the proposed method are compared to when traditional measurements are incorporated into the DSSE calculation using a sample-and-hold technique.

One more paragraph saying something like: “In summary, the research presented in this dissertation takes a step towards more effectively monitoring and controlling active distribution

networks. The contributions to this field will lead to more efficient energy use, reducing the cost and carbon footprint of future power networks.”

Publication

- 1- A Review on State Estimation Techniques in Active Distribution Networks: Existing Practices and Their Challenges, (Published in Sustainability Journal), February 2022.
- 2- Simultaneous State Estimation and Bad Data Detection on PMU Measurements in Active Distribution Power Networks, (Published in NAPS), October 2022.
- 3- Random Forest Meta Learner for Generating Pseudo-Measurements in Active Distribution Power Networks, (Published in ISGT USA), January 2023.
- 4- Distribution System State Estimation and False Data Injection Attack Detection with a Multi-Output Deep Neural Network, (Published in in Energies Journal). February 2023.
- 5- Distribution System State Estimation Using Hybrid Traditional and Advanced Measurements for Grid Modernization, (Published in Applied Science Journal), April 2023.
- 6- Incorporating Temporal and Meteorological Data for Generating Pseudo-measurements in Active Distribution Power Networks, (Published in CAI USA), June 2023.
- 7- An Overview of Supervised Machine Learning Approaches for Applications in Active Distribution Networks, (Published in Energies Journal). July 2023.

REFERENCES CITED

- [1] S. S. M. and O. T. J. Duncan, Glover J., *Power System Analysis and Design*. 2012.
- [2] F. Li *et al.*, “Smart transmission grid: Vision and framework,” *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 168–177, 2010, doi: 10.1109/TSG.2010.2053726.
- [3] D. Deka, V. Kekatos, and G. Cavraro, “Learning Distribution Grid Topologies: A Tutorial,” *IEEE Trans. Smart Grid*, vol. PP, p. 1, 2022, doi: 10.1109/TSG.2023.3271902.
- [4] L. Gomes, A. Coelho, and Z. Vale, “Assessment of Energy Customer Perception, Willingness, and Acceptance to Participate in Smart Grids—A Portuguese Survey,” *Energies*, vol. 16, no. 1, pp. 1–16, 2023, doi: 10.3390/en16010270.
- [5] A. Aghahosseini *et al.*, “Energy system transition pathways to meet the global electricity demand for ambitious climate targets and cost competitiveness,” *Appl. Energy*, vol. 331, no. November 2022, 2023, doi: 10.1016/j.apenergy.2022.120401.
- [6] P. B *et al.*, “Deep Learning for Intelligent Demand Response and Smart Grids: A Comprehensive Survey,” pp. 1–25, 2021, [Online]. Available: <http://arxiv.org/abs/2101.08013>.
- [7] Z. R. S. C. Palanisamy, Sivaraman, “Introduction to Smart Power Systems,” in *Artificial Intelligence-based Smart Power Systems*, 2023, pp. 1–13.
- [8] S. N. Khan, S. A. A. Kazmi, A. Altamimi, Z. A. Khan, and M. A. Alghassab, “Smart Distribution Mechanisms—Part I: From the Perspectives of Planning,” *Sustain.*, vol. 14, no. 23, 2022, doi: 10.3390/su142316308.
- [9] G. Dileep, “A survey on smart grid technologies and applications,” *Renew. Energy*, vol. 146, pp. 2589–2625, 2020, doi: 10.1016/j.renene.2019.08.092.
- [10] J. Fan and S. Borlase, “The evolution of distribution,” *IEEE Power Energy Mag.*, vol. 7, no. 2, pp. 63–68, 2009, doi: 10.1109/MPE.2008.931392.
- [11] S. A. A. Kazmi, M. K. Shahzad, A. Z. Khan, and D. R. Shin, *Smart distribution networks: A review of modern distribution concepts from a planning perspective*, vol. 10, no. 4. 2017.
- [12] C. Muscas, M. Pau, P. Pegoraro, and S. Sulis, “Smart electric energy measurements in power distribution grids,” *IEEE Instrum. Meas. Mag.*, vol. 18, no. 1, pp. 17–21, 2015, doi: 10.1109/MIM.2015.7016676.
- [13] B. R. Pokhrel, B. Bak-Jensen, and J. R. Pillai, “Integrated approach for network observability and state estimation in active distribution grid,” *Energies*, vol. 12, no. 11, 2019, doi: 10.3390/en12112230.

- [14] F. Ahmad, A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş, “Distribution system state estimation-A step towards smart grid,” *Renew. Sustain. Energy Rev.*, vol. 81, no. November 2016, pp. 2659–2671, 2018, doi: 10.1016/j.rser.2017.06.071.
- [15] B. Stephen, “Machine Learning Applications in Power System Condition Monitoring,” *Energies*, vol. 15, no. 5, 2022, doi: 10.3390/en15051808.
- [16] J. Zhao *et al.*, “Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work,” *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, 2019, doi: 10.1109/TPWRS.2019.2894769.
- [17] A. Bose, “Smart transmission grid applications and their supporting infrastructure,” *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 11–19, 2010, doi: 10.1109/TSG.2010.2044899.
- [18] G. T. Heydt, “The next generation of power distribution systems,” *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 225–235, 2010, doi: 10.1109/TSG.2010.2080328.
- [19] M. Yilmaz and P. T. Krein, “Review of the impact of vehicle-to-grid technologies on distribution systems and utility interfaces,” *IEEE Trans. Power Electron.*, vol. 28, no. 12, pp. 5673–5689, 2013, doi: 10.1109/TPEL.2012.2227500.
- [20] W. M. Lin, J. H. Teng, and S. J. Chen, “A highly efficient algorithm in treating current measurements for the branch-current-based distribution state estimation,” *IEEE Trans. Power Deliv.*, vol. 16, no. 3, pp. 433–439, 2001, doi: 10.1109/61.924823.
- [21] T. Niknam and B. B. Firouzi, “A practical algorithm for distribution state estimation including renewable energy sources,” *Renew. Energy*, vol. 34, no. 11, pp. 2309–2316, 2009, doi: 10.1016/j.renene.2009.03.005.
- [22] D. Thukaram, J. Jerome, and C. Surapong, “Robust three-phase state estimation algorithm for distribution networks,” *Electr. Power Syst. Res.*, vol. 55, no. 3, pp. 191–200, 2000, doi: 10.1016/S0378-7796(00)00088-2.
- [23] I. Roytelman and S. M. Shahidehpour, “State Estimation for Electric Power Distribution Systems in Quasi real-Time Conditions,” *IEEE Trans. Power Deliv.*, vol. 8, no. 4, pp. 2009–2015, 1993, doi: 10.1109/61.248315.
- [24] A. W. Kelley, “State e.,” vol. 9, no. 3, pp. 1601–1609, 1994.
- [25] Y. Wang, Q. Chen, T. Hong, and C. Kang, “Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges,” *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3125–3148, 2019, doi: 10.1109/TSG.2018.2818167.
- [26] J. Zheng, D. W. Gao, and L. Lin, “Smart meters in smart grid: An overview,” *IEEE Green Technol. Conf.*, pp. 57–64, 2013, doi: 10.1109/GreenTech.2013.17.

- [27] R. Singh, B. C. Pal, and R. A. Jabr, "Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement," *IET Gener. Transm. Distrib.*, vol. 4, no. 1, pp. 50–59, 2010, doi: 10.1049/iet-gtd.2009.0167.
- [28] A. Angioni, T. Schlösser, F. Ponci, and A. Monti, "Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, 2016, doi: 10.1109/TIM.2015.2454673.
- [29] M. Pau, P. A. Pegoraro, and S. Sulis, "Performance of three-phase WLS distribution system state estimation approaches," *2015 IEEE Int. Work. Appl. Meas. Power Syst. AMPS 2015 - Proc.*, pp. 138–143, 2015, doi: 10.1109/AMPS.2015.7312752.
- [30] Z. Jia, J. Chen, and Y. Liao, "State estimation in distribution system considering effects of AMI data," *Conf. Proc. - IEEE SOUTHEASTCON*, 2013, doi: 10.1109/SECON.2013.6567406.
- [31] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaahedi, "Distribution System State Estimation Based on Nonsynchronized Smart Meters," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2919–2928, 2015, doi: 10.1109/TSG.2015.2429640.
- [32] M. Pau, P. A. Pegoraro, and S. Sulis, "Performance of three-phase WLS distribution system state estimation approaches," *2015 IEEE Int. Work. Appl. Meas. Power Syst. AMPS 2015 - Proc.*, 2015, doi: 10.1109/AMPS.2015.7312752.
- [33] C. H. Oh, S. Il Go, J. H. Choi, S. J. Ahn, and S. Y. Yun, "Voltage estimation method for power distribution networks using high-precision measurements," *Energies*, vol. 13, no. 9, 2020, doi: 10.3390/en13092385.
- [34] G. Paramo, A. Bretas, and S. Meyn, "Research Trends and Applications of PMUs," *Energies*, vol. 15, no. 15, pp. 1–32, 2022, doi: 10.3390/en15155329.
- [35] C. Zhang, Y. Jia, Z. Xu, L. L. Lai, and K. P. Wong, "Optimal PMU placement considering state estimation uncertainty and voltage controllability," *IET Gener. Transm. Distrib.*, vol. 11, no. 18, pp. 4465–4475, 2017, doi: 10.1049/iet-gtd.2016.2023.
- [36] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *arXiv*, vol. 10, no. 2, pp. 2312–2322, 2018.
- [37] J. Kim, H. T. Kim, and S. Choi, "Performance Criterion of Phasor Measurement Units for Distribution System State Estimation," *IEEE Access*, vol. 7, pp. 106372–106384, 2019, doi: 10.1109/ACCESS.2019.2932940.
- [38] J. Xu, Y. Jin, T. Zheng, and G. Meng, "On State Estimation Modeling of Smart Distribution Networks: A Technical Review," *Energies*, vol. 16, no. 4, 2023, doi: 10.3390/en16041891.

- [39] Sonal and D. Ghosh, “Novel trends in resilience assessment of a distribution system using synchrophasor application: A literature review,” *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 8, 2021, doi: 10.1002/2050-7038.12934.
- [40] S. Chakrabarti, E. Kyriakides, G. Ledwich, and A. Ghosh, “Inclusion of PMU current phasor measurements in a power system state estimator,” *IET Gener. Transm. Distrib.*, vol. 4, no. 10, pp. 1104–1115, 2010, doi: 10.1049/iet-gtd.2009.0398.
- [41] J. Romero Agüero, E. Takayesu, D. Novosel, and R. Masiello, “Grid modernization: challenges and opportunities,” *Electr. J.*, vol. 30, no. 4, pp. 1–6, 2017, doi: 10.1016/j.tej.2017.03.008.
- [42] G. Frigo and F. Grasso Toro, “On-line performance assessment for improved sensor data aggregation in power system metrology,” *Meas. Sensors*, vol. 18, 2021, doi: 10.1016/j.measen.2021.100186.
- [43] D. Novosel and V. Centeno, “Benefits of Synchronized-Measurement Technology for Power-Grid Applications,” pp. 1–9, 2007.
- [44] J. Lin *et al.*, “Design , Implementation and Application of Synchrophasor-Based Online State Estimation in a Large-Scale Real Power Grid,” 2021, doi: 10.17775/CSEEJPES.2021.08550.
- [45] M. Asprou, S. Chakrabarti, and E. Kyriakides, “A two-stage state estimator for dynamic monitoring of power systems,” *IEEE Syst. J.*, vol. 11, no. 3, pp. 1767–1776, 2017, doi: 10.1109/JSYST.2014.2375951.
- [46] N. M. Manousakis and G. N. Korres, “A hybrid power system state estimator using synchronized and unsynchronized sensors,” *Int. Trans. Electr. Energy Syst.*, vol. 28, no. 8, pp. 1–19, 2018, doi: 10.1002/etep.2580.
- [47] O. Darmis and G. Korres, “A Survey on Hybrid SCADA/WAMS State Estimation Methodologies in Electric Power Transmission Systems,” *Energies*, vol. 16, no. 2, 2023, doi: 10.3390/en16020618.
- [48] S. A. Shezan *et al.*, “Evaluation of Different Optimization Techniques and Control Strategies of Hybrid Microgrid: A Review,” *Energies*, vol. 16, no. 4, pp. 1–30, 2023, doi: 10.3390/en16041792.
- [49] K. Nainar and F. Iov, “Smart meter measurement-based state estimation for monitoring of low-voltage distribution grids,” *Energies*, vol. 13, no. 20, 2020, doi: 10.3390/en13205367.
- [50] U. Kuhar, M. Pantos, G. Kosec, and A. Svirgelj, “The Impact of Model and Measurement Uncertainties on a State Estimation in Three-Phase Distribution Networks,” *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3301–3310, 2019, doi: 10.1109/TSG.2018.2823398.

- [51] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "False data injection threats in active distribution systems: A comprehensive survey," *Futur. Gener. Comput. Syst.*, vol. 140, pp. 344–364, 2022, doi: 10.1016/j.future.2022.10.021.
- [52] M. Kamal, M. Farajollahi, H. Nazaripouya, and H. Mohsenian-Rad, "Cyberattacks against Event-Based Analysis in Micro-PMUs: Attack Models and Counter Measures," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1577–1588, 2021, doi: 10.1109/TSG.2020.3029937.
- [53] R. Deng, P. Zhuang, and H. Liang, "False Data Injection Attacks Against State Estimation in Power Distribution Systems," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2871–2881, 2019, doi: 10.1109/TSG.2018.2813280.
- [54] P. Zhuang, R. Deng, and H. Liang, "False Data Injection Attacks against State Estimation in Multiphase and Unbalanced Smart Distribution Systems," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6000–6013, 2019, doi: 10.1109/TSG.2019.2895306.
- [55] G. Cheng, Y. Lin, J. Yan, J. Zhao, and L. Bai, "Model-Measurement Data Integrity Attacks," *IEEE Trans. Smart Grid*, vol. PP, p. 1, 2023, doi: 10.1109/TSG.2023.3253781.
- [56] V. Asefi, S., Mitrovic, M., Ćetenović, D., Levi, V., Gryazina, E., & Terzija, "Anomaly detection and classification in power system state estimation: Combining model-based and data-driven methods," *Sustain. Energy, Grids Networks*, vol. 35, p. 101116, 2023.
- [57] Y. F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 33–43, 2012, doi: 10.1109/MSP.2012.2187037.
- [58] A. Angioni, J. Shang, F. Ponci, and A. Monti, "Real-Time Monitoring of Distribution System Based on State Estimation," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 10, pp. 2234–2243, 2016, doi: 10.1109/TIM.2016.2583239.
- [59] H. T. Reda, A. Anwar, A. N. Mahmood, and Z. Tari, "A Taxonomy of Cyber Defence Strategies Against False Data Attacks in Smart Grids," *ACM Comput. Surv.*, vol. 55, no. 14, 2023, doi: 10.1145/3592797.
- [60] S. M. Miraftebadeh, F. Foiadelli, M. Longo, and M. Pasetti, "A Survey of Machine Learning Applications for Power System Analytics," *Proc. - 2019 IEEE Int. Conf. Environ. Electr. Eng. 2019 IEEE Ind. Commer. Power Syst. Eur. IEEEIC/I CPS Eur. 2019*, 2019, doi: 10.1109/IEEEIC.2019.8783340.
- [61] H. Yang, X. Liu, D. Zhang, T. Chen, C. Li, and W. Huang, "Machine learning for power system protection and control," *Electr. J.*, vol. 34, no. 1, p. 106881, 2021, doi: 10.1016/j.tej.2020.106881.

- [62] H. Wang, C. Ma, and L. Zhou, "A brief review of machine learning and its application," *Proc. - 2009 Int. Conf. Inf. Eng. Comput. Sci. ICIECS 2009*, pp. 1–4, 2009, doi: 10.1109/ICIECS.2009.5362936.
- [63] J. Alzubi, A. Nayyar, and A. Kumar, "Machine Learning from Theory to Algorithms: An Overview," *J. Phys. Conf. Ser.*, vol. 1142, no. 1, 2018, doi: 10.1088/1742-6596/1142/1/012012.
- [64] M. Batta, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 18, no. 8, pp. 381–386, 2018, doi: 10.21275/ART20203995.
- [65] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," vol. 349, no. 6245, 2015.
- [66] J. H. Menke, N. Bornhorst, and M. Braun, "Distribution system monitoring for smart power grids with distributed generation using artificial neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 113, no. April, pp. 472–480, 2019, doi: 10.1016/j.ijepes.2019.05.057.
- [67] N. Neranjan Thilakarathne, K. K. Mohan, L. Surekha, and A. Hussain, "Smart Grid: A Survey of Architectural Elements, Machine Learning and Deep Learning Applications and Future Directions," *J. Intell. Syst. Internet Things*, pp. 32–42, 2021, doi: 10.54216/jisiot.030103.
- [68] A. S. Musleh, G. Chen, and Z. Y. Dong, "A Survey on the Detection Algorithms for False Data Injection Attacks in Smart Grids," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2218–2234, 2020, doi: 10.1109/TSG.2019.2949998.
- [69] M. Mohammadpourfard, A. Sami, and A. Seifi, "A statistical unsupervised method against false data injection attacks: A visualization-based approach," *Expert Syst. Appl.*, vol. 84, pp. 242–261, 2017, doi: 10.1016/j.eswa.2017.05.013.
- [70] C. Tu, X. He, Z. Shuai, and F. Jiang, "Big data issues in smart grid – A review," *Renew. Sustain. Energy Rev.*, vol. 79, no. July 2016, pp. 1099–1107, 2017, doi: 10.1016/j.rser.2017.05.134.
- [71] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of Big Data and Machine Learning in Smart Grid, and Associated Security Concerns: A Review," *IEEE Access*, vol. 7, pp. 13960–13988, 2019, doi: 10.1109/ACCESS.2019.2894819.
- [72] V. K. Ponnusamy *et al.*, "A comprehensive review on sustainable aspects of big data analytics for the smart grid," *Sustain.*, vol. 13, no. 23, pp. 1–35, 2021, doi: 10.3390/su132313322.
- [73] O. H. Abu-Rub, A. Y. Fard, M. F. Umar, M. Hosseinzadehtaher, and M. B. Shadmands, "Towards intelligent power electronics-dominated grid via machine learning techniques," *IEEE Power Electron. Mag.*, vol. 8, no. 1, pp. 28–38, 2021, doi:

10.1109/MPEL.2020.3047506.

- [74] F. Aminifar, M. Abedini, T. Amraee, P. Jafarian, M. H. Samimi, and M. Shahidehpour, *A review of power system protection and asset management with machine learning techniques*, vol. 13, no. 4. Springer Berlin Heidelberg, 2022.
- [75] A. Ahmed and M. Khalid, “A review on the selected applications of forecasting models in renewable power systems,” *Renew. Sustain. Energy Rev.*, vol. 100, no. March 2018, pp. 9–21, 2019, doi: 10.1016/j.rser.2018.09.046.
- [76] J. Hwang, D. Suh, and M. O. Otto, “Forecasting electricity consumption in commercial buildings using a machine learning approach,” *Energies*, vol. 13, no. 22, 2020, doi: 10.3390/en13225885.
- [77] S. Ray, “A Quick Review of Machine Learning Algorithms,” *Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Perspectives Prospect. Com. 2019*, pp. 35–39, 2019, doi: 10.1109/COMITCon.2019.8862451.
- [78] S. Raschka, “Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning,” 2018, [Online]. Available: <http://arxiv.org/abs/1811.12808>.
- [79] H. Salehi and R. Burgueño, “Emerging artificial intelligence methods in structural engineering,” *Eng. Struct.*, vol. 171, no. November 2017, pp. 170–189, 2018, doi: 10.1016/j.engstruct.2018.05.084.
- [80] V. Nasteski, “An overview of the supervised machine learning methods,” *Horizons.B*, vol. 4, no. December 2017, pp. 51–62, 2017, doi: 10.20544/horizons.b.04.1.17.p05.
- [81] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, *A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science*. 2020.
- [82] J. E. van Engelen and H. H. Hoos, “A survey on semi-supervised learning,” *Mach. Learn.*, vol. 109, no. 2, pp. 373–440, 2020, doi: 10.1007/s10994-019-05855-6.
- [83] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, and D. Hassabis, “Reinforcement Learning, Fast and Slow,” *Trends Cogn. Sci.*, vol. 23, no. 5, pp. 408–422, 2019, doi: 10.1016/j.tics.2019.02.006.
- [84] P. Y. Glorionnec, “Reinforcement Learning: an Overview,” no. September, pp. 17–35, 2000, [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.9.4135&rep=rep1&type=pdf>.
- [85] R. Salman and V. Kecman, “Regression as classification,” *Conf. Proc. - IEEE SOUTHEASTCON*, pp. 1–6, 2012, doi: 10.1109/SECon.2012.6196887.

- [86] X. Su, X. Yan, and C. L. Tsai, "Linear regression," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 4, no. 3, pp. 275–294, 2012, doi: 10.1002/wics.1198.
- [87] S. I. Bangdiwala, "Regression: simple linear," *Int. J. Inj. Contr. Saf. Promot.*, vol. 25, no. 1, pp. 113–115, 2018, doi: 10.1080/17457300.2018.1426702.
- [88] Alexopoulos EC, "Introduction to Multivariate Regression Analysis," *Hippokratia*, vol. 14, no. 1, pp. 23–28, 2005, [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049417/pdf/hippokratia-14-23.pdf>
[http://as.osa.org/abstract.cfm?URI=as-64-12-335A%5CnC:/Vasili/Laboratory/Articles/studyplasma.Data/PDF//2010 Omenetto review-3017268501/2010 Omenetto review.pdf](http://as.osa.org/abstract.cfm?URI=as-64-12-335A%5CnC:/Vasili/Laboratory/Articles/studyplasma.Data/PDF//2010%20Omenetto%20review-3017268501/2010%20Omenetto%20review.pdf).
- [89] B. Smola, A. J., & Schölkopf, "A tutorial on support vector regression," *J. Neuroeng. Rehabil.*, vol. 14, no. 1, pp. 199–222, 2002, doi: 10.1186/s12984-021-00806-6.
- [90] D. F. Specht, "A General Regression Neural Network," *IEEE Trans. Neural Networks*, vol. 2, no. 6, pp. 568–576, 1991, doi: 10.1109/IJCNN.2007.4371258.
- [91] D. M. J. Tax and R. P. W. Duin, "Using two-class classifiers for multiclass classification," *Proc. - Int. Conf. Pattern Recognit.*, vol. 16, no. 2, pp. 124–127, 2002, doi: 10.1109/icpr.2002.1048253.
- [92] G. Tsoumakas and I. Katakis, "Multi-Label Classification: An Overview," *Database Technol.*, pp. 309–319, 2011, doi: 10.4018/978-1-60566-058-5.ch021.
- [93] M. Maalouf, "Logistic regression in data analysis: An overview," *Int. J. Data Anal. Tech. Strateg.*, vol. 3, no. 3, pp. 281–299, 2011, doi: 10.1504/IJDATS.2011.041335.
- [94] G. Wang, "A survey on training algorithms for support vector machine classifiers," *Proc. - 4th Int. Conf. Networked Comput. Adv. Inf. Manag. NCM 2008*, vol. 1, pp. 123–128, 2008, doi: 10.1109/NCM.2008.103.
- [95] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, 2021, doi: 10.38094/jastt20165.
- [96] F. Féraud, R., & Clérot, "A methodology to explain neural network classification," *Neural networks*, vol. 15, pp. 237–246, 2002, [Online]. Available: http://ac.elsa-cdn.com/S0893608001001277/1-s2.0-S0893608001001277-main.pdf?_tid=014dc778-07fe-11e7-bef2-00000aabb0f02&acdnat=1489417498_65d78a487ce5847b13d2f7030ab219a4.
- [97] P. Sedgwick, "Simple linear regression," *BMJ*, vol. 346, no. 7904, pp. 1–2, 2013, doi: 10.1136/bmj.f2340.

- [98] R. M. Heiberger and E. Neuwirth, *RExcel: R Through Excel*. 2009.
- [99] P. Sinha, "Multivariate Polynomial Regression in Data Mining: Methodology, Problems and Solutions," *Int. J. Sci. Eng. Res.*, vol. 4, no. 12, pp. 962–965, 2013, [Online]. Available: <http://www.ijser.org>.
- [100] A. Demaris, "A Tutorial in Logistic Regression Published by : National Council on Family Relations," *J. Marriage Fam.*, vol. 57, no. 4, pp. 956–968, 1995.
- [101] A. J. Scott, D. W. Hosmer, and S. Lemeshow, "Applied Logistic Regression.," *Biometrics*, vol. 47, no. 4, p. 1632, 1991, doi: 10.2307/2532419.
- [102] A. Pradhan, "SUPPORT VECTOR MACHINE-A Survey," vol. 2, no. 8, pp. 82–85, 2012.
- [103] R. Awad, M., Khanna, R., Awad, M., & Khanna, *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*. 2015.
- [104] P. Anand, R. Rastogi, and S. Chandra, "A class of new Support Vector Regression models," *Appl. Soft Comput. J.*, vol. 94, p. 106446, 2020, doi: 10.1016/j.asoc.2020.106446.
- [105] J. M. Moguerza and A. Muñoz, "Support vector machines with applications," *Stat. Sci.*, vol. 21, no. 3, pp. 322–336, 2006, doi: 10.1214/088342306000000493.
- [106] S. Ghosh, A. Dasgupta, and A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification," *Proc. Int. Conf. Intell. Sustain. Syst. ICISS 2019*, no. Iciss, pp. 24–28, 2019, doi: 10.1109/ISS1.2019.8908018.
- [107] K. M. Lin and C. J. Lin, "A study of reduced support vector machines," *IEEE Trans. Neural Networks*, vol. 14, no. 6, pp. 1449–1459, 2003, doi: 10.1109/TNN.2003.820828.
- [108] R. J. Lewis, D. Ph, and W. C. Street, "An Introduction to Classification and Regression Tree (CART) Analysis," *2000 Annu. Meet. Soc. Acad. Emerg. Med.*, no. 310, p. 14p, 2000, [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4103&rep=rep1&type=pdf>.
- [109] J. Singh Kushwah, A. Kumar, S. Patel, R. Soni, A. Gawande, and S. Gupta, "Comparative study of regressor and classifier with decision tree using modern tools," *Mater. Today Proc.*, vol. 56, pp. 3571–3576, 2022, doi: 10.1016/j.matpr.2021.11.635.
- [110] B. de Ville, "Decision trees," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 5, no. 6, pp. 448–455, 2013, doi: 10.1002/wics.1278.
- [111] M. N. Anyanwu and S. G. Shiva, "Comparative Analysis of Serial Decision Tree Classification Algorithms," *Int. J. Comput. Sci. Secur.*, vol. 3, no. 3, pp. 230–240, 2009, [Online]. Available:

<http://www.cscjournals.org/csc/manuscriptinfo.php?ManuscriptCode=72.73.66.82.82.44.55.49.99>.

- [112] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, “An introduction to decision tree modeling,” *J. Chemom.*, vol. 18, no. 6, pp. 275–285, 2004, doi: 10.1002/cem.873.
- [113] K. Suzuki, *Artificial Neural Networks - Methodological Advanced and Biomedical Applications*, no. August. 2017.
- [114] P. Marius-Constantin, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, “Multilayer perceptron and neural networks,” *WSEAS Trans. Circuits Syst.*, vol. 8, no. 7, pp. 579–588, 2009.
- [115] A. Canziani, A. Paszke, and E. Culurciello, “An Analysis of Deep Neural Network Models for Practical Applications,” pp. 1–7, 2016, [Online]. Available: <http://arxiv.org/abs/1605.07678>.
- [116] S. Bianco, R. Cadene, L. Celona, and P. Napoletano, “Benchmark analysis of representative deep neural network architectures,” *IEEE Access*, vol. 6, pp. 64270–64277, 2018, doi: 10.1109/ACCESS.2018.2877890.
- [117] V. M. Sonde, P. P. Shirpurkar, M. S. Giripunje, and P. P. Ashtankar, *Deep learning techniques: an overview*, vol. 1141. 2021.
- [118] F. Hoffmann, T. Bertram, R. Mikut, M. Reischl, and O. Nelles, “Benchmarking in classification and regression,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 5, pp. 1–17, 2019, doi: 10.1002/widm.1318.
- [119] M. N. Hossin, M., & Sulaiman, “A review on evaluation metrics for data classification evaluations,” vol. 5, no. 2, pp. 1–11, 2015.
- [120] S. K. Tiwary, J. Pal, and C. K. Chanda, “Multiple-Classification of Power System States Using Multidimensional Neural Network,” *J. Inst. Eng. Ser. B*, no. 0123456789, pp. 1–8, 2023, doi: 10.1007/s40031-023-00892-1.
- [121] V. Plevris, G. Solorzano, N. P. Bakas, and M. E. A. Ben Seghier, “Investigation of Performance Metrics in Regression Analysis and Machine Learning-Based Prediction Models,” *World Congr. Comput. Mech. ECCOMAS Congr.*, pp. 0–25, 2022, doi: 10.23967/eccomas.2022.155.
- [122] D. Della Giustina, M. Pau, P. A. Pegoraro, F. Ponci, and S. Sulis, “Electrical distribution system state estimation: Measurement issues and challenges,” *IEEE Instrum. Meas. Mag.*, vol. 17, no. 6, pp. 36–42, 2014, doi: 10.1109/MIM.2014.6968929.
- [123] A. E. Saldaña-González, A. Sumper, M. Aragüés-Peñalba, and M. Smolnikar, “Advanced

- distribution measurement technologies and data applications for smart grids: A review,” *Energies*, vol. 13, no. 14, pp. 1–34, 2020, doi: 10.3390/en13143730.
- [124] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, “Effects of measurements and pseudomeasurements correlation in distribution system state estimation,” *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 2813–2823, 2014, doi: 10.1109/TIM.2014.2318391.
 - [125] J. W. Taylor, “Short-term electricity demand forecasting using double seasonal exponential smoothing,” *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, 2003, doi: 10.1057/palgrave.jors.2601589.
 - [126] P. Su *et al.*, “Recent trends in load forecasting technology for the operation optimization of distributed energy system,” *Energies*, vol. 10, no. 9, 2017, doi: 10.3390/en10091303.
 - [127] R. Jain and V. Mahajan, “Load forecasting and risk assessment for energy market with renewable based distributed generation,” *Renew. Energy Focus*, vol. 42, pp. 190–205, 2022, doi: 10.1016/j.ref.2022.06.007.
 - [128] H. S. Hippert, C. E. Pedreira, and R. C. Souza, “Neural networks for short-term load forecasting: A review and evaluation,” *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, 2001, doi: 10.1109/59.910780.
 - [129] A. Khotanzad, R. C. Hwang, A. Abaye, and D. Maratukulam, “An Adaptive Modular Artificial Neural Network Hourly Load Forecaster and its Implementation at Electric Utilities,” *IEEE Trans. Power Syst.*, vol. 10, no. 3, pp. 1716–1722, 1995, doi: 10.1109/59.466468.
 - [130] O. A. Omitaomu and H. Niu, “Artificial Intelligence Techniques in Smart Grid: A Survey,” *Smart Cities*, vol. 4, no. 2, pp. 548–568, 2021, doi: 10.3390/smartsities4020029.
 - [131] E. Manitsas, R. Singh, B. C. Pal, and S. Member, “Distribution System State Estimation Using an Artificial Neural Network Approach for Pseudo Measurement Modeling,” vol. 27, no. 4, pp. 1888–1896, 2012.
 - [132] Z. Cao, Y. Wang, C. C. Chu, and R. Gadh, “Robust pseudo-measurement modeling for three-phase distribution systems state estimation,” *Electr. Power Syst. Res.*, vol. 180, no. May 2019, p. 106138, 2020, doi: 10.1016/j.epsr.2019.106138.
 - [133] K. Dehghanpour, Y. Yuan, Z. Wang, and F. Bu, “A Game-Theoretic Data-Driven Approach for Pseudo-Measurement Generation in Distribution System State Estimation,” *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 5942–5951, 2019, doi: 10.1109/TSG.2019.2893818.
 - [134] M. E. Tipping, “The relevance vector machine,” *Adv. Neural Inf. Process. Syst.*, no. x, pp. 653–658, 2000.
 - [135] B. P. Hayes, J. K. Gruber, and M. Prodanovic, “A Closed-Loop State Estimation Tool for

- MV Network Monitoring and Operation,” *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2116–2125, 2015, doi: 10.1109/TSG.2014.2378035.
- [136] K. Dehghanpour, Y. Yuan, Z. Wang, and F. Bu, “A game-theoretic data-driven approach for pseudo-measurement generation in distribution system state estimation,” *arXiv*, vol. 10, no. 6, pp. 5942–5951, 2018.
 - [137] Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, “A Multi-Timescale Data-Driven Approach to Enhance Distribution System Observability,” *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3168–3177, 2019, doi: 10.1109/TPWRS.2019.2893821.
 - [138] G. Hong and Y. S. Kim, “Supervised Learning Approach for State Estimation of Unmeasured Points of Distribution Network,” *IEEE Access*, vol. 8, pp. 113918–113931, 2020, doi: 10.1109/ACCESS.2020.3003049.
 - [139] L. Pasic, A. Pasic, B. Hartmann, and I. Vokony, “The Application of Artificial Neural Networks to Pseudo Measurement Modeling in Distribution Networks,” *2021 IEEE Madrid PowerTech, PowerTech 2021 - Conf. Proc.*, 2021, doi: 10.1109/PowerTech46648.2021.9494844.
 - [140] Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, “A deep learning method for short-term residential load forecasting in smart grid,” *IEEE Access*, vol. 8, pp. 55785–55797, 2020, doi: 10.1109/ACCESS.2020.2981817.
 - [141] G. Hafeez, K. S. Alimgeer, and I. Khan, “Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid,” *Appl. Energy*, vol. 269, no. March, p. 114915, 2020, doi: 10.1016/j.apenergy.2020.114915.
 - [142] S. Wu and Y. Li, “State Estimation of Distribution Network Considering Data Compatibility,” *Energy Power Eng.*, vol. 12, no. 04, pp. 73–83, 2020, doi: 10.4236/epe.2020.124b008.
 - [143] M. Göl, S. Member, and A. Abur, “A Hybrid State Estimator For Systems With Limited Number of PMUs,” *IEEE Trans. power Syst. Syst.*, vol. 30, no. 3, pp. 1511–1517, 2015.
 - [144] L. Ramesh, S. P. Chowdhury, S. Chowdhury, and N. Chakraborty, “Ring main distribution state estimation through optimal meter placement,” *IET Conf. Publ.*, vol. 2012, no. 624 CP, pp. 341–348, 2012, doi: 10.1049/cp.2012.2236.
 - [145] F. Marini and B. Walczak, “Particle swarm optimization (PSO). A tutorial,” *Chemom. Intell. Lab. Syst.*, vol. 149, pp. 153–165, 2015, doi: 10.1016/j.chemolab.2015.08.020.
 - [146] S. Wang and G. Liu, “State estimation for active distribution networks basedon multisource hybrid measurements,” *China Int. Conf. Electr. Distrib. CICED*, vol. 2016-Septe, no. Ciced, pp. 10–13, 2016, doi: 10.1109/CICED.2016.7575943.

- [147] M. Mao, J. Xu, Z. Wu, Q. Hu, and X. Dou, "A Multiarea State Estimation for Distribution Networks under Mixed Measurement Environment," *IEEE Trans. Ind. Informatics*, vol. 18, no. 6, pp. 3620–3629, 2022, doi: 10.1109/TII.2021.3119949.
- [148] X. Kong *et al.*, "A hybrid state estimator based on SCADA and PMU measurements for medium voltage distribution system," *Appl. Sci.*, vol. 8, no. 9, 2018, doi: 10.3390/app8091527.
- [149] R. Z. S. Santos and J. R. C. Orillaza, "Distribution system state estimator using SCADA and μ PMU measurements," *Asia-Pacific Power Energy Eng. Conf. APPEEC*, vol. 2018-Octob, pp. 469–474, 2018, doi: 10.1109/APPEEC.2018.8566281.
- [150] M. Huang, Z. Wei, G. Sun, and H. Zang, "Hybrid State Estimation for Distribution Systems With AMI and SCADA Measurements," *IEEE Access*, vol. 7, pp. 120350–120359, 2019, doi: 10.1109/access.2019.2937096.
- [151] C. S. Kumar, K. Rajawat, S. Chakrabarti, and B. C. Pal, "Robust distribution system state estimation with hybrid measurements," *IET Gener. Transm. Distrib.*, vol. 14, no. 16, pp. 3250–3259, 2020, doi: 10.1049/iet-gtd.2020.0260.
- [152] S. Shanmugapriya and D. Maharajan, "A fast Broyden's approximation-based weighted least square state estimation for power systems," *Int. J. Numer. Model. Electron. Networks, Devices Fields*, vol. 34, no. 2, pp. 1–13, 2021, doi: 10.1002/jnm.2831.
- [153] Y. Liu, Y. Wang, and Q. Yang, "Spatio-Temporal Generative Adversarial Network based Power Distribution Network State Estimation with Multiple Time-scale Measurements," *IEEE Trans. Ind. Informatics*, vol. PP, no. 9, pp. 1–8, 2023, doi: 10.1109/tii.2023.3234624.
- [154] X. He, C. Li, M. Du, H. Dong, and P. Li, "Hybrid Measurements-Based Fast State Estimation for Power Distribution System," *IEEE Access*, vol. 9, pp. 21112–21122, 2021, doi: 10.1109/ACCESS.2021.3055223.
- [155] M. Farajollahi, A. Shahsavari, and H. Mohsenian-Rad, "Tracking State Estimation in Distribution Networks Using Distribution-level Synchrophasor Data," *IEEE Power Energy Soc. Gen. Meet.*, vol. 2018-Augus, pp. 1–5, 2018, doi: 10.1109/PESGM.2018.8586273.
- [156] F. Ahmad, A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş, "Distribution system state estimation-A step towards smart grid," *Renew. Sustain. Energy Rev.*, vol. 81, no. November 2016, pp. 2659–2671, 2018, doi: 10.1016/j.rser.2017.06.071.
- [157] J. Watitwa and K. Awodele, "A Review on Active Distribution System State Estimation," *Proc. - 2019 South. African Univ. Power Eng. Conf. Mechatronics/Pattern Recognit. Assoc. South Africa, SAUPEC/RobMech/PRASA 2019*, pp. 726–731, 2019, doi: 10.1109/RoboMech.2019.8704833.
- [158] M. S. N. G. Sarada Devi and G. Yesuratnam, *Comparison of State Estimation Process on*

Transmission and Distribution Systems. Springer International Publishing, 2020.

- [159] A. F. Sanchez Cifuentes, “Enhancing the Observability of Distribution Systems State Estimation,” p. 51, 2019, [Online]. Available: <https://fsu.digital.flvc.org/islandora/object/fsu%3A722569#tabs-1>.
- [160] B. Hayes and M. Prodanović, “State estimation techniques for electric power distribution systems,” *Proc. - UKSim-AMSS 8th Eur. Model. Symp. Comput. Model. Simulation, EMS 2014*, pp. 303–308, 2014, doi: 10.1109/EMS.2014.76.
- [161] M. Majdoub, A. Belfqih, J. Boukherouaa, O. Sabri, B. Cheddadi, and T. Haidi, “A review on distribution system state estimation techniques,” *Proc. 2018 6th Int. Renew. Sustain. Energy Conf. IRSEC 2018*, pp. 1–6, 2018, doi: 10.1109/IRSEC.2018.8702952.
- [162] G. Wang, G. B. Giannakis, J. Chen, and J. Sun, “Distribution system state estimation : an overview,” *Front. Inf. Technol. Electron.*, vol. 20, no. 1, pp. 4–17, 2019, [Online]. Available: [//doi.org/10.1631/FITEE.1800590](https://doi.org/10.1631/FITEE.1800590).
- [163] S. Radhoush, T. Vannoy, B. M. Whitaker, and H. Nehrir, “Random Forest Meta Learner for Generating Pseudo-Measurements in Active Distribution Power Networks,” *2023 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf.*, pp. 1–5, 2023, doi: 10.1109/ISGT51731.2023.10066389.
- [164] H. Liao and J. V. Milanović, “Pathway to cost-efficient state estimation of future distribution networks,” *IEEE Power Energy Soc. Gen. Meet.*, vol. 2016-Novem, 2016, doi: 10.1109/PESGM.2016.7741337.
- [165] I. Džafić, M. Gilles, R. A. Jabr, B. C. Pal, and S. Henselmeyer, “Real time estimation of loads in radial and unsymmetrical three-phase distribution networks,” *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4839–4848, 2013, doi: 10.1109/TPWRS.2013.2277888.
- [166] S. Radhoush, M. Bahramipanah, H. Nehrir, and Z. Shahooei, “A Review on State Estimation Techniques in Active Distribution Networks: Existing Practices and Their Challenges,” *Sustainability*, vol. 14, no. 5, p. 2520, 2022, doi: 10.3390/su14052520.
- [167] M. Fotopoulou, S. Petridis, I. Karachalios, and D. Rakopoulos, “A Review on Distribution System State Estimation Algorithms,” *Appl. Sci.*, vol. 12, no. 21, pp. 1–21, 2022, doi: 10.3390/app122111073.
- [168] W. Commerell, “The Grid of the Future,” *Green Energy Technol.*, no. april, pp. 239–269, 2020, doi: 10.1007/978-3-030-41952-3_10.
- [169] A. Ehsan and Q. Yang, “State-of-the-art techniques for modelling of uncertainties in active distribution network planning: A review,” *Appl. Energy*, vol. 239, no. November 2018, pp. 1509–1523, 2019, doi: 10.1016/j.apenergy.2019.01.211.

- [170] T. Wuest, D. Weimer, C. Irgens, and K. D. Thoben, "Machine learning in manufacturing: Advantages, challenges, and applications," *Prod. Manuf. Res.*, vol. 4, no. 1, pp. 23–45, 2016, doi: 10.1080/21693277.2016.1192517.
- [171] G. L. Rajora, M. A. Sanz-Bobi, and C. M. Domingo, "Application of Machine Learning Methods for Asset Management on Power Distribution Networks," *Emerg. Sci. J.*, vol. 6, no. 4, pp. 905–920, 2022, doi: 10.28991/ESJ-2022-06-04-017.
- [172] B. Azimian, R. Sen 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, 2022, doi: 10.1109/TIM.2022.3167722.
- [173] J. H. Menke, N. Bornhorst, and M. Braun, "Distribution system monitoring for smart power grids with distributed generation using artificial neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 113, no. July 2018, pp. 472–480, 2019, doi: 10.1016/j.ijepes.2019.05.057.
- [174] M. Q. Tran, A. S. Zamzam, and P. H. Nguyen, "Enhancement of Distribution System State Estimation Using Pruned Physics-Aware Neural Networks," *2021 IEEE Madrid PowerTech, PowerTech 2021 - Conf. Proc.*, 2021, doi: 10.1109/PowerTech46648.2021.9494950.
- [175] B. Zargar, F. Ponci, and A. Monti, "Evaluation of Computational Complexity for Distribution Systems State Estimation," *IEEE Trans. Instrum. Meas.*, vol. 72, no. 1557, 2023, doi: 10.1109/TIM.2023.3260268.
- [176] M. Andryscio, A. Nötzli, F. Brown, R. Jhala, and D. Stefan, "Towards verified, constant-time floating point operations," *Proc. ACM Conf. Comput. Commun. Secur.*, pp. 1369–1382, 2018, doi: 10.1145/3243734.3243766.
- [177] P. A. Dinda, "Online prediction of the running time of tasks," *IEEE Int. Symp. High Perform. Distrib. Comput. Proc.*, pp. 383–394, 2001, doi: 10.1145/378420.378836.
- [178] B. Zargar, A. Angioni, F. Ponci, and A. Monti, "Multiarea Parallel Data-Driven Three-Phase Distribution System State Estimation Using Synchrophasor Measurements," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6186–6202, 2020, doi: 10.1109/TIM.2020.2967512.
- [179] S. Chen, Z. Wei, G. Sun, N. Lu, Y. Sun, and Y. Zhu, "Multi-area distributed three-phase state estimation for unbalanced active distribution networks," *J. Mod. Power Syst. Clean Energy*, vol. 5, no. 5, pp. 767–776, 2017, doi: 10.1007/s40565-016-0237-0.
- [180] C. Muscas, M. Pau, P. A. Pegoraro, S. Sulis, F. Ponci, and A. Monti, "Multiarea distribution system state estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 5, pp. 1140–1148, 2015, doi: 10.1109/TIM.2014.2365406.

- [181] M. Ferdowsi, A. Benigni, A. Lowen, B. Zargar, A. Monti, and F. Ponci, "A scalable data-driven monitoring approach for distribution systems," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 5, pp. 1292–1305, 2015, doi: 10.1109/TIM.2015.2398991.
- [182] X. B. Jin, R. J. R. Jeremiah, T. L. Su, Y. T. Bai, and J. L. Kong, "The new trend of state estimation: From model-driven to hybrid-driven methods," *Sensors*, vol. 21, no. 6, pp. 1–25, 2021, doi: 10.3390/s21062085.
- [183] 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, 2019, doi: 10.1109/TPWRS.2019.2919157.
- [184] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, "Data-Driven Learning-Based Optimization for Distribution System State Estimation," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4796–4805, 2019, doi: 10.1109/TPWRS.2019.2909150.
- [185] X. Li, A. Scaglione, and T. H. Chang, "A framework for phasor measurement placement in hybrid state estimation via Gauss-Newton," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 824–832, 2014, doi: 10.1109/TPWRS.2013.2283079.
- [186] W. Wang, N. Yu, F. Rahmatian, and S. Pandey, "Where to Install Distribution Phasor Measurement Units to Obtain Accurate State Estimation Results?," *IEEE Power Energy Soc. Gen. Meet.*, vol. 2022-July, pp. 1–5, 2022, doi: 10.1109/PESGM48719.2022.9916656.
- [187] F. Ahmad, M. Tariq, and A. Farooq, "A novel ANN-based distribution network state estimator," *Int. J. Electr. Power Energy Syst.*, vol. 107, no. January 2018, pp. 200–212, 2019, doi: 10.1016/j.ijepes.2018.11.019.
- [188] Y. Ma, F. Liu, X. Zhou, and Z. Gao, "Key technologies of smart distribution grid," *2016 IEEE Int. Conf. Mechatronics Autom. IEEE ICMA 2016*, pp. 2639–2643, 2016, doi: 10.1109/ICMA.2016.7558983.
- [189] A. W. Kelley, "State estimation for real-time monitoring of distribution systems," vol. 9, no. 3, pp. 1601–1609, 1994.
- [190] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid - The new and improved power grid: A survey," *IEEE Commun. Surv. Tutorials*, vol. 14, no. 4, pp. 944–980, 2012, doi: 10.1109/SURV.2011.101911.00087.
- [191] S. R. Salkuti, "Challenges, issues and opportunities for the development of smart grid," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 2, pp. 1179–1186, 2020, doi: 10.11591/ijece.v10i2.pp1179-1186.
- [192] X. Liu and Z. Li, "False data attack models, impact analyses and defense strategies in the electricity grid," *Electr. J.*, vol. 30, no. 4, pp. 35–42, 2017, doi: 10.1016/j.tej.2017.04.001.

- [193] A. Sayghe *et al.*, “Survey of machine learning methods for detecting false data injection attacks in power systems,” *IET Smart Grid*, vol. 3, no. 5, pp. 581–595, 2020, doi: 10.1049/iet-stg.2020.0015.
- [194] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, 2011, doi: 10.1145/1952982.1952995.
- [195] H. T. Reda, A. Anwar, and A. Mahmood, “Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts,” *Renew. Sustain. Energy Rev.*, vol. 163, no. February, p. 112423, 2022, doi: 10.1016/j.rser.2022.112423.
- [196] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, “False data injection on state estimation in power systems-attacks, impacts, and defense: A survey,” *IEEE Trans. Ind. Informatics*, vol. 13, no. 2, pp. 411–423, 2017, doi: 10.1109/TII.2016.2614396.
- [197] Y. Huang *et al.*, “Bad data injection in smart grid: Attack and defense mechanisms,” *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 27–33, 2013, doi: 10.1109/MCOM.2013.6400435.
- [198] R. Z. S. Santos and J. R. C. Orillaza, “Distribution system state estimator using SCADA and μ PMU measurements: An FDI attack vulnerability analysis,” *Asia-Pacific Power Energy Eng. Conf. APPEEC*, vol. 2018-Octob, pp. 469–474, 2018, doi: 10.1109/APPEEC.2018.8566281.
- [199] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. Vincent Poor, “Distributed models for sparse attack construction and state vector estimation in the smart grid,” *2012 IEEE 3rd Int. Conf. Smart Grid Commun. SmartGridComm 2012*, pp. 306–311, 2012, doi: 10.1109/SmartGridComm.2012.6486001.
- [200] Y. Wang, A. Xing, Z. Qu, X. Han, H. Dong, and P. M. Georgievitch, “False data injection attack detection based on interval affine state estimation,” *Electr. Power Syst. Res.*, vol. 210, no. November 2021, p. 108100, 2022, doi: 10.1016/j.epsr.2022.108100.
- [201] M. Mohammadpourfard, Y. Weng, and M. Tajdinian, “Benchmark of machine learning algorithms on capturing future distribution network anomalies,” *IET Gener. Transm. Distrib.*, vol. 13, no. 8, pp. 1441–1455, 2019, doi: 10.1049/iet-gtd.2018.6801.
- [202] N. Bhusal, M. Gautam, and M. Benidris, “Detection of Cyber Attacks on Voltage Regulation in Distribution Systems Using Machine Learning,” *IEEE Access*, vol. 9, pp. 40402–40416, 2021, doi: 10.1109/ACCESS.2021.3064689.
- [203] M. Ganjkhani, M. Gilanifar, J. Giraldo, and M. Parvania, “Integrated Cyber and Physical Anomaly Location and Classification in Power Distribution Systems,” *IEEE Trans. Ind. Informatics*, vol. 17, no. 10, pp. 7040–7049, 2021, doi: 10.1109/TII.2021.3065080.
- [204] E. Naderi, A. Aydeger, and A. Asrari, “Detection of False Data Injection Cyberattacks

- Targeting Smart Transmission/Distribution Networks,” *2022 IEEE Conf. Technol. Sustain. SusTech 2022*, pp. 224–229, 2022, doi: 10.1109/SusTech53338.2022.9794237.
- [205] S. Sapna, “Backpropagation Learning Algorithm Based on Levenberg Marquardt Algorithm,” pp. 393–398, 2012, doi: 10.5121/csit.2012.2438.
- [206] M. Kamal, A. Shahsavari, and H. Mohsenian-Rad, “Poisoning Attack against Event Classification in Distribution Synchrophasor Measurements,” *2021 IEEE Int. Conf. Commun. Control. Comput. Technol. Smart Grids, SmartGridComm 2021*, pp. 327–332, 2021, doi: 10.1109/SmartGridComm51999.2021.9632336.
- [207] F. L. Grando, A. E. Lazzaretti, and M. Moreto, “The Impact of PMU Data Precision and Accuracy on Event Classification in Distribution Systems,” *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1372–1382, 2022, doi: 10.1109/TSG.2021.3126268.
- [208] T. Seidl, “Nearest Neighbor Classification,” *Encycl. Database Syst.*, vol. 1, pp. 1885–1890, 2009, doi: 10.1007/978-0-387-39940-9_561.
- [209] Y. Karaca and C. Cattani, “Naive Bayesian classifier,” *Comput. Methods Data Anal.*, pp. 229–250, 2018, doi: 10.1515/9783110496369-007.
- [210] T. G. Dietterich and Oregon, “Ensemble methods in machine learning. In: International Workshop on Multiple Classifier Models,” *Oncogene*, vol. 12, no. 2, pp. 1–15, 2000.
- [211] Y. Yang, W. Hong, and S. Li, “Deep ensemble learning based probabilistic load forecasting in smart grids,” *Energy*, vol. 189, p. 116324, 2019, doi: 10.1016/j.energy.2019.116324.
- [212] P. Liashchynskyi and P. Liashchynskyi, “Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS,” no. 2017, pp. 1–11, 2019, [Online]. Available: <http://arxiv.org/abs/1912.06059>.
- [213] Dr. Joseph A. Shaw, “ORSL Weather Station.” the Montana State University weather station.
- [214] S. Dubey, J. N. Sarvaiya, and B. Seshadri, “Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world - A review,” *Energy Procedia*, vol. 33, pp. 311–321, 2013, doi: 10.1016/j.egypro.2013.05.072.
- [215] P. et Al, “Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825–2830, 2011.” .
- [216] P. F. Smith, S. Ganesh, and P. Liu, “A comparison of random forest regression and multiple linear regression for prediction in neuroscience,” *J. Neurosci. Methods*, vol. 220, no. 1, pp. 85–91, 2013, doi: 10.1016/j.jneumeth.2013.08.024.
- [217] O. of E. E. & R. E. (EERE), “Commercial and Residential Hourly Load Profiles.”

<https://openei.org/datasets/files/961/pub/>.

- [218] M. Sedighizadeh, M. Jahangir, M. Gandomkar, and S. Esfandeh, “Distributed generation location and capacity effect on voltage stability of distribution network,” *Int. Conf. Math. Methods Comput. Tech. Electr. Eng. - Proc.*, vol. 25, no. c, pp. 89–94, 2010.
- [219] N. R. E. L. PVWatts version 5 manual, “<https://pvwatts.nrel.gov/>.”.
- [220] J. Chen *et al.*, “Impact analysis of false data injection attacks on power system static security assessment,” *J. Mod. Power Syst. Clean Energy*, vol. 4, no. 3, pp. 496–505, 2016, doi: 10.1007/s40565-016-0223-6.
- [221] A. Monticelli and A. Garcia, “Reliable Bad Data Processing for Real-Time State Estimation,” *IEEE Power Eng. Rev.*, vol. PER-3, no. 5, pp. 31–32, 1983, doi: 10.1109/MPER.1983.5519146.
- [222] G. Hug and J. A. Giampapa, “Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks,” *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012, doi: 10.1109/TSG.2012.2195338.
- [223] S. Chowdhury, S. Chowdhury, and P. Crossley, *Microgrids and Active Distribution Networks, IET RENEWABLE ENERGY SERIES 6*. 2009.
- [224] M. Mobtahej, K. Esapour, S. Z. Tajalli, and M. Mohammadi, “Effective demand response and GANs for optimal constraint unit commitment in solar-tidal based microgrids,” *IET Renew. Power Gener.*, vol. 16, no. 16, pp. 3485–3495, 2022, doi: 10.1049/rpg2.12331.
- [225] Z. Wang, B. Zhang, M. Mobtahej, A. Baziar, and B. Khan, “Advanced Reactive Power Compensation of Wind Power Plant Using PMU Data,” *IEEE Access*, vol. 9, pp. 67006–67014, 2021, doi: 10.1109/ACCESS.2021.3075966.
- [226] Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao, “On false data-injection attacks against power system state estimation: Modeling and countermeasures,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 717–729, 2014, doi: 10.1109/TPDS.2013.92.
- [227] M. Abdelmalak, V. Venkataramanan, and R. MacWan, “A Survey of Cyber-Physical Power System Modeling Methods for Future Energy Systems,” *IEEE Access*, vol. 10, no. August, pp. 99875–99896, 2022, doi: 10.1109/ACCESS.2022.3206830.
- [228] S. Barja-Martinez, M. Aragüés-Peñalba, Í. Munné-Collado, P. Lloret-Gallego, E. Bullich-Massagué, and R. Villafafila-Robles, “Artificial intelligence techniques for enabling Big Data services in distribution networks: A review,” *Renew. Sustain. Energy Rev.*, vol. 150, p. 111459, 2021, doi: 10.1016/j.rser.2021.111459.
- [229] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, “Data-driven learning-based optimization for distribution system state estimation,” *arXiv*, vol. 34, no. 6, pp. 4796–4805, 2018.

- [230] C. Rudin *et al.*, “Machine learning for the New York City power grid,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 328–345, 2012, doi: 10.1109/TPAMI.2011.108.
- [231] M. Abdelmalak, S. Member, H. Hooshyar, and S. Member, “Real-Time EMT-Phasor Co-Simulation Modeling for Large-Scale Power Grids : Challenges and Solutions,” pp. 1–5, 2022.
- [232] B. R. N. Anderson, M. Ieee, A. Boulanger, W. B. Powell, and W. Scott, “Adaptive Stochastic Control for the Smart Grid,” *Proc. IEEE*, vol. 99, no. August, pp. 14–21, 2011.
- [233] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor, “Machine Learning Methods for Attack Detection in the Smart Grid,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 8, pp. 1773–1786, 2016, doi: 10.1109/TNNLS.2015.2404803.
- [234] S. Radhoush, T. Vannoy, B. M. Whitaker, and H. Nehrir, “Simultaneous State Estimation and Bad Data Detection on PMU Measurements in Active Distribution Power Networks,” pp. 1–6, 2023, doi: 10.1109/naps56150.2022.10012150.
- [235] and A. G. E. Abur, Ali, *Power system state estimation: theory and implementation*. CRC press, 2004.
- [236] T. Ahmad, H. Chen, Y. Guo, and J. Wang, “A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review,” *Energy Build.*, vol. 165, pp. 301–320, 2018, doi: 10.1016/j.enbuild.2018.01.017.
- [237] A. G. Abur, A., & Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.
- [238] Y. Gao and N. Yu, “State estimation for unbalanced electric power distribution systems using AMI data,” *2017 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. ISGT 2017*, 2017, doi: 10.1109/ISGT.2017.8085999.
- [239] C. Muscas, S. Sulis, A. Angioni, F. Ponci, and A. Monti, “Impact of different uncertainty sources on a three-phase state estimator for distribution networks,” *IEEE Trans. Instrum. Meas.*, vol. 63, no. 9, pp. 2200–2209, 2014, doi: 10.1109/TIM.2014.2308352.
- [240] A. Majumdar and B. C. Pal, “A three-phase state estimation in unbalanced distribution networks with switch modelling,” *2016 IEEE 1st Int. Conf. Control. Meas. Instrumentation, C. 2016*, no. Cmi, pp. 474–478, 2016, doi: 10.1109/CMI.2016.7413793.
- [241] X. Kong, Y. Chen, C. Yong, X. Ma, and J. Kong, “Stepwise robust distribution system state estimation considering PMU measurement,” *J. Renew. Sustain. Energy*, vol. 11, no. 2, 2019, doi: 10.1063/1.5064532.
- [242] O. Simeone, “An Introduction to Quantum Machine Learning for Engineers,” *Found. Trends Signal Process.*, vol. 16, no. 1–2, pp. 1–223, 2022, doi: 10.1561/2000000118.

- [243] D. Zhang, X. Han, and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," *CSEE J. Power Energy Syst.*, vol. 4, no. 3, pp. 362–370, 2018, doi: 10.17775/cseejpes.2018.00520.
- [244] A. Moradzadeh, M. Mohammadpourfard, C. Konstantinou, I. Genc, T. Kim, and B. Mohammadi-Ivatloo, "Electric load forecasting under False Data Injection Attacks using deep learning," *Energy Reports*, vol. 8, pp. 9933–9945, 2022, doi: 10.1016/j.egyr.2022.08.004.
- [245] K. Y. Lee, J. S. Park, and Y. S. Kim, "Optimal placement of pmu to enhance supervised learning-based pseudo-measurement modelling accuracy in distribution network," *Energies*, vol. 14, no. 22, 2021, doi: 10.3390/en14227767.
- [246] J. Vijaychandra, B. R. V. Prasad, V. K. Darapureddi, B. V. Rao, and Ł. Knypiński, "A Review of Distribution System State Estimation Methods and Their Applications in Power Systems," *Electronics*, vol. 12, no. 3, p. 603, 2023, doi: 10.3390/electronics12030603.
- [247] D. K. Panda and S. Das, "Smart grid architecture model for control, optimization and data analytics of future power networks with more renewable energy," *J. Clean. Prod.*, vol. 301, p. 126877, 2021, doi: 10.1016/j.jclepro.2021.126877.
- [248] Y. Kabalci, "A survey on smart metering and smart grid communication," *Renew. Sustain. Energy Rev.*, vol. 57, pp. 302–318, 2016, doi: 10.1016/j.rser.2015.12.114.
- [249] R. Morello, S. C. Mukhopadhyay, Z. Liu, D. Slomovitz, and S. R. Samantaray, "Advances on sensing technologies for smart cities and power grids: A review," *IEEE Sens. J.*, vol. 17, no. 23, pp. 7596–7610, 2017, doi: 10.1109/JSEN.2017.2735539.
- [250] B. Yildiz, J. I. Bilbao, J. Dore, and A. B. Sproul, "Recent advances in the analysis of residential electricity consumption and applications of smart meter data," *Appl. Energy*, vol. 208, no. January, pp. 402–427, 2017, doi: 10.1016/j.apenergy.2017.10.014.
- [251] R. A. Walling, R. Saint, R. C. Dugan, J. Burke, and L. A. Kojovic, "Summary of distributed resources impact on power delivery systems," *IEEE Trans. Power Deliv.*, vol. 23, no. 3, pp. 1636–1644, 2008, doi: 10.1109/TPWRD.2007.909115.
- [252] N. Uribe-Pérez, L. Hernández, D. de la Vega, and I. Angulo, "State of the Art and Trends Review of Smart Metering in Electricity Grids," *Appl. Sci.*, vol. 6, no. 3, pp. 1–24, 2016, doi: 10.3390/app6030068.
- [253] U.S. Department of Energy Office of Electricity Delivery and Energy Reliability (DOE-OE), "Factors Affecting PMU Installation Costs," no. September, p. 12, 2014.
- [254] J. Singh, M. Kumar, A. Sharma, G. Pandey, K. Chae, and S. Lee, "The Optimal Operation of Active Distribution Networks with Smart Systems," *Intech*, vol. 11, no. tourism, p. 13, 2016, [Online]. Available: <https://www.intechopen.com/books/advanced-biometric->

technologies/liveness-detection-in-biometrics.

- [255] V. Terzija *et al.*, “Wide-area monitoring, protection, and control of future electric power networks,” *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, 2011, doi: 10.1109/JPROC.2010.2060450.
- [256] A. Gómez-Expósito, A. Abur, P. Rousseaux, A. De La Villa Jaén, and C. Gómez-Quiles, “On the use of PMUs in power system state estimation,” *17th Power Syst. Comput. Conf. PSCC 2011*, 2011.
- [257] Y. Liu, J. Li, and L. Wu, “State Estimation of Three-Phase Four-Conductor Distribution Systems With Real-Time Data From Selective Smart Meters,” *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2632–2643, 2019, doi: 10.1109/TPWRS.2019.2892726.
- [258] Ke Li, “State estimation for power distribution system and measurement impacts,” *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 911–916, May 1996, doi: 10.1109/59.496174.
- [259] G. Tang, Y. Han, J. Wang, and Q. Zhao, “Smart Meter Data Analytics for Distribution Network,” *Chinese Control Conf. CCC*, vol. 2018-July, no. 4, pp. 8882–8887, 2018, doi: 10.23919/ChiCC.2018.8483503.
- [260] F. F. Wu, “Power system state estimation: a survey,” *Int. J. Electr. Power Energy Syst.*, vol. 12, no. 2, pp. 80–87, 1990, doi: 10.1016/0142-0615(90)90003-T.
- [261] M. Pau, P. A. Pegoraro, and S. Sulis, “WLS distribution system state estimator based on voltages or branch-currents: Accuracy and performance comparison,” *Conf. Rec. - IEEE Instrum. Meas. Technol. Conf.*, pp. 493–498, 2013, doi: 10.1109/I2MTC.2013.6555467.
- [262] H. Wang, W. Zhang, and Y. Liu, “A Robust Measurement Placement Method for Active Distribution System State Estimation Considering Network Reconfiguratio,” *2108 IEEE Trans. SMART GRID*, vol. 9, no. 3, pp. 2108–2117, 2018.
- [263] H. Wang and N. N. Schulz, “A Revised Branch Current-Based Distribution System State Estimation Algorithm and Meter Placement Impact,” *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 207–213, 2004, doi: 10.1109/TPWRS.2003.821426.
- [264] M. Pau, P. A. Pegoraro, A. Monti, C. Muscas, F. Ponci, and S. Sulis, “Impact of Current and Power Measurements on Distribution System State Estimation Uncertainty,” *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3992–4002, 2019, doi: 10.1109/TIM.2018.2883844.
- [265] C. Technologies and C. C. License, “Three-Phase State Estimation for Distribution-Grid Analytics,” 2021.
- [266] S. Chakrabarti, E. Kyriakides, G. Ledwich, and A. Ghosh, “A comparative study of the methods of inclusion of pmu current phasor measurements in a hybrid state estimator,”

IEEE PES Gen. Meet. PES 2010, pp. 1–7, 2010, doi: 10.1109/PES.2010.5589582.

- [267] R. Hoffman, “Practical state estimation for electric distribution networks,” *2006 IEEE PES Power Syst. Conf. Expo. PSCE 2006 - Proc.*, pp. 510–517, 2006, doi: 10.1109/PSCE.2006.296366.
- [268] C. Hawk and A. Kaushiva, “Cybersecurity and the Smarter Grid,” *Electr. J.*, vol. 27, no. 8, pp. 84–95, 2014, doi: 10.1016/j.tej.2014.08.008.
- [269] M. Kabiri, S. Member, N. Amjady, and S. Member, “A New Hybrid State Estimation Considering Different Accuracy Levels of PMU and SCADA Measurements,” *IEEE Trans. Instrum. Meas.*, vol. PP, no. 9, pp. 1–12, 2018.
- [270] G. N. Korres and N. M. Manousakis, “A state estimation algorithm for monitoring topology changes in distribution systems,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–8, 2012, doi: 10.1109/PESGM.2012.6345126.
- [271] H. Borchani, G. Varando, C. Bielza, and B. Monte, “A survey on multi-output regression,” pp. 1–27.
- [272] M. Pau, F. Ponci, A. Monti, S. Sulis, C. Muscas, and P. A. Pegoraro, “An Efficient and Accurate Solution for Distribution System State Estimation with Multiarea Architecture,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 5, pp. 910–919, 2017, doi: 10.1109/TIM.2016.2642598.
- [273] S. Kayalvizhi and V. K. Vinod, “Optimal planning of active distribution networks with hybrid distributed energy resources using grid-based multi-objective harmony search algorithm,” *Appl. Soft Comput. J.*, vol. 67, pp. 387–398, 2018, doi: 10.1016/j.asoc.2018.03.009.
- [274] B. Zargar, A. Angioni, F. Ponci, and A. Monti, “Multiarea Parallel Data-Driven Three-Phase Distribution System State Estimation Using Synchrophasor Measurements,” *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6186–6202, 2020, doi: 10.1109/TIM.2020.2967512.