energies

MDPI

Article

Distribution System State Estimation and False Data Injection Attack Detection with a Multi-Output Deep Neural Network

Sepideh Radhoush, Trevor Vannoy, Kaveen Liyanage, Bradley M. Whitaker and Hashem Nehrir *

Electrical and Computer engineering Department, Montana State University, Bozeman, MT 59717, USA
* Correspondence: hnehrir@montana.edu

Abstract: Distribution system state estimation (DSSE) has been introduced to monitor distribution grids; however, due to the incorporation of distributed generations (DGs), traditional DSSE methods are not able to reveal the operational conditions of active distribution networks (ADNs). DSSE calculation depends heavily on real measurements from measurement devices in distribution networks. However, the accuracy of real measurements and DSSE results can be significantly affected by false data injection attacks (FDIAs). Conventional FDIA detection techniques are often unable to identify FDIAs into measurement data. In this study, a novel deep neural network approach is proposed to simultaneously perform DSSE calculation (i.e., regression) and FDIA detection (i.e., binary classification) using real measurements. In the proposed work, the classification nodes in the DNN allow us to identify which measurements on which phasor measurement unit (PMU), if any, were affected. In the proposed approach, we aim to show that the proposed method can perform DSSE calculation and identify FDIAs from the available measurements simultaneously with high accuracy. We compare our proposed method to the traditional approach of detecting FDIAs and performing SE calculations separately; moreover, DSSE results are compared with the weighted least square (WLS) algorithm, which is a common model-based method. The proposed method achieves better DSSE performance than the WLS method and the separate DSSE/FDIA method in presence of erroneous measurements; our method also executes faster than the other methods. The effectiveness of the proposed method is validated using two FDIA schemes in two case studies: one using a modified IEEE 33-bus distribution system without DGs, and the other using a modified IEEE 69-bus system with DGs. The results illustrated that the accuracy and *F*1-score of the proposed method are better than when performing binary classification only. The proposed method successfully detected the FDIAs on each PMU measurement. Moreover, the results of DSSE calculation from the proposed method has a better performance compared to the regression-only method, and the WLS methods in the presence of bad data.

Keywords: distribution system state estimation; false data injection attacks; deep neural network; weighted least square; active distribution network; bad data detection

check for updates

Citation: Radhoush, S.; Vannoy, T.; Liyanage, K.; Whitaker, B.M.; Nehrir, H. Distribution System State Estimation and False Data Injection Attack Detection with a Multi-Output Deep Neural Network. *Energies* **2023**, *16*, 2288. https://doi.org/10.3390/en16052288

Academic Editors: Theofilos A. Papadopoulos and Eleftherios O. Kontis

Received: 5 February 2023
Revised: 20 February 2023
Accepted: 24 February 2023
Published: 27 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The concept of false data injection attacks (FDIAs) in power systems was first studied in [1]; additional studies about the damaging effects and threats of these new attacks quickly followed [2,3]. Because of the transition from traditional power networks to smart grids, more smart devices and communication infrastructures are required to enable the reliable and efficient performance of smart grids [4]. However, despite the progress in the power network structure, attackers attempt to disrupt the performance of power networks by manipulating the data obtained from measurement devices [5]. The goals of attackers are to falsify the data and measurement information in power networks. Therefore, state estimation (SE) results can be influenced by FDIAs due to dependencies of state estimation calculations on measurement information and network topology [6]. SE results are submitted to the control center for further processing including optimal

power flow, contingency analysis, security analysis, etc. 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 [7–9].

The goal of BDD is to determine the existence of erroneous bad data on measurement information. Traditional bad data detection (BDD) schemes compute the ℓ_2-norm of the residual measurement after SE calculations are performed [10,11]. 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 [12]. State estimation techniques, attacks, and defense strategies on transmission networks have been well established [13,14]. Unfortunately, these approaches cannot be applied in distribution networks (DNs) due to their differences with transmission networks [15]. Some features of DNs include:

1. Unlike transmission networks, which have a mesh structure, DNs generally have radial or weakly meshed configuration.
2. DNs generally consist of many buses compared to transmission networks, making installation of measurement devices at all buses in DNs economically impractical.
3. DNs normally have large resistance to reactance ratios (r/x), compared to transmission lines.

In addition to these differences, renewable energy sources are becoming more common in DNs [16–18]. These sources typically introduce higher variance and inconsistency, making it more difficult to perform state estimation. Because of these reasons, new SE methods should be developed that consider the characteristics of DNs: multiple renewable energy sources [19], electric vehicles, variable loads, etc. Moreover, the integration of different technologies and components in active distribution networks (ADNs) must emphasize the security aspects of these networks, including the ability to detect cyberattacks such as FDIAs [20,21].

In the last decade, machine learning approaches have been widely used and developed for control and monitoring in power networks [22–25]. In [26], 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 applied widely to identify malicious data injection. Faster execution time and accurate results are two main advantages of using machine learning approaches over conventional BDD methods [27–29]. 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.

Different machine learning methods, including supervised [30–32], semi-supervised [33,34], and unsupervised learning [35,36], are used to explore the detection of FDIAs in many different fields. In [37], a machine learning approach is utilized to identify cyberattacks such as structured query language injection attack (SQLIA. In [38], a new machine learning method is proposed to identify false data injection attack on an information of technology (IOT) system. In [39], a supervised-machine learning algorithm is used to classify different failure parts of a wind turbine.

In this study, we focus our machine learning efforts in FDIA detection in power distribution networks. FDIA detection is considered a supervised binary classification problem. In [40], 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 [41], 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 distribution networks. 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 distribution grid. In [42], two distributed sparse state estimation and attack detection methods are studied to make a DN observable and to perform FDIA detection locally in distribution networks. In [43], the affine interval state estimation method is applied to identify attacks

in measurement data by considering the upper and lower boundaries of the state variables. In [44], a new method is proposed to identify types of faults and cyberattack locations simultaneously by utilizing a deep neural network (DNN) method. The authors call their method fault and attack location and classification (FALCON); this method is categorized as a multi-output classification approach. In [45], a new method is proposed to identify the presence of corrupted measurement data and the location of compromised micro-PMUs in order to ignore the corrupted measurement devices as a defense strategy. In [46], two separate DNN models are designed to perform DSSE and topology detection from available synchronized measurements for unbalanced ADNs.

In 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, multi-output classification problems are solved by a single DNN model. In contrast, this paper uses a single DNN model for simultaneously performing DSSE and FDIA detection using PMU measurements as inputs. 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. 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 paper are summarized as follows:

1. We design a single DNN model to simultaneously perform DSSE calculation and FDIA detection based on PMU measurement inputs. 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 calculation is 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. We show that the proposed method accurately calculates state estimation variables, even in the presence of erroneous measurements.
7. The execution time comparison between the proposed method, the disjoint DNN model for DSSE calculation and FDIA detection, and the model-based approach is calculated. The results indicate that performing FDIA detection and DSSE calculation simultaneously lead to a significant decrease in execution time.

2. Power System State Estimation

State estimation calculations are essential for continually improving the performance and management of power networks [47]. Different state estimation techniques have been developed and used for transmission networks for several years, but these methods cannot be used at the distribution level directly due to differences between transmission and distribution networks. The distribution system state estimation (DSSE) enhances monitoring and controlling of distribution grids effectively and efficiently. Moreover, state estimation results can be utilized for load forecasting, stability analysis, optimal power flow, bad data detection, and energy market analysis [48,49]. The weighted least square (WLS) method is one of the traditional methods which is effective in both distribution and transmission grids.

WLS Formulation for State Estimation

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

$$z = h(x) + e_z, \tag{1}$$

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 $h(x)$. The state variable vector is given by $x = [x_1 \dots x_N]^T$ 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 of zero mean dimension and be a Gaussian random variable with covariance matrix $R_z = diag\{\sigma_{z_1}^2, \sigma_{z_2}^2, \cdots, \sigma_{z_m}^2\}$. For instance, the power flow equations can be expressed as:

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

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

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}, respectively; 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:

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

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

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

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

In this formulation, w_i is the weight associated with the ith measurement, and M is the number of available measurements to perform the 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:

$$J(x) = [z - h(x)]^T W [z - h(x)]. \tag{7}$$

$W_{M \times M}$ 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 $J(x)$; Reference [47] In IGN, the following function is solved at each iteration k:

$$G(x_k) \Delta x_k = H_k^T W [z - h(x_k)], \tag{8}$$

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 used to calculate the new state as follows:

$$x_{k+1} = x_k + \Delta x_k. \tag{9}$$

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 [50,51]. 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 coordinates as $x = [\delta_2, \cdots, \delta_N, V_1, \cdots, V_N]$, 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 distribution grid. The voltage magnitude is 1 p.u. and 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 $x = [\delta_{1\varnothing}, \ldots, \delta_{N\varnothing}, V_{1\varnothing}, \ldots, V_{N\varnothing}]$ where the phase angle $\delta_{1\varnothing}$ is not zero any more [52–54].

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 distribution grids is altered. Second, the Jacobian matrix has a different structure [55].

3. False Data Injection Attacks (FDIAs)

To evaluate the operating status of power networks, including voltage magnitude and phase angle of buses, state estimation is made on the basis of available measurements. Unfortunately, the state variables can be manipulated by injecting FDIAs into meter measurements [56], 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:

$$z_a = z + a. \tag{10}$$

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:

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

The general effect of FDIAs on measurements and DSSE procedures are shown in Figure 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. The SE results are then sent to the control and management center for further processing, including bad data detection using appropriate methods.

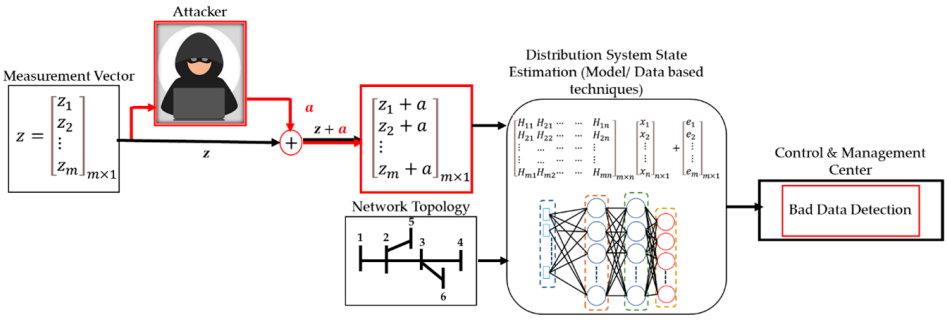

Figure 1. General Effect of FDIAs on Measurements and DSEE Procedure.

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 [57]. 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:

$$y_i = \begin{cases} 0, & if\ a = 0 \\ 1, & if\ a \neq 0 \end{cases} \tag{12}$$

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.

A 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 [58,59]; however, this technique requires more time and data for a training phase [60,61]. 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 2.

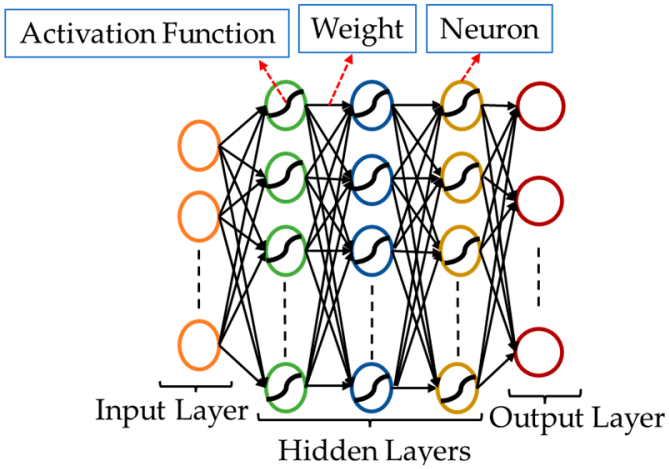

Figure 2. DNN Configuration.

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).

4. 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 3a,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 3c, FDIA detection and DSSE calculation are performed simultaneously using a single DNN model. The method is compared to traditional approaches that perform BDD and DSSE calculation using two independent DNN models. Traditional approaches use a regression-based DNN to perform DSSE calculation and a classification-based DNN to detect FDIAs. A description of regression-based DNNs and classification-based DNNs in the context of DSSE and BDD is provided in [62]. In this paper, we assume 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 $z_a = z + a$.)

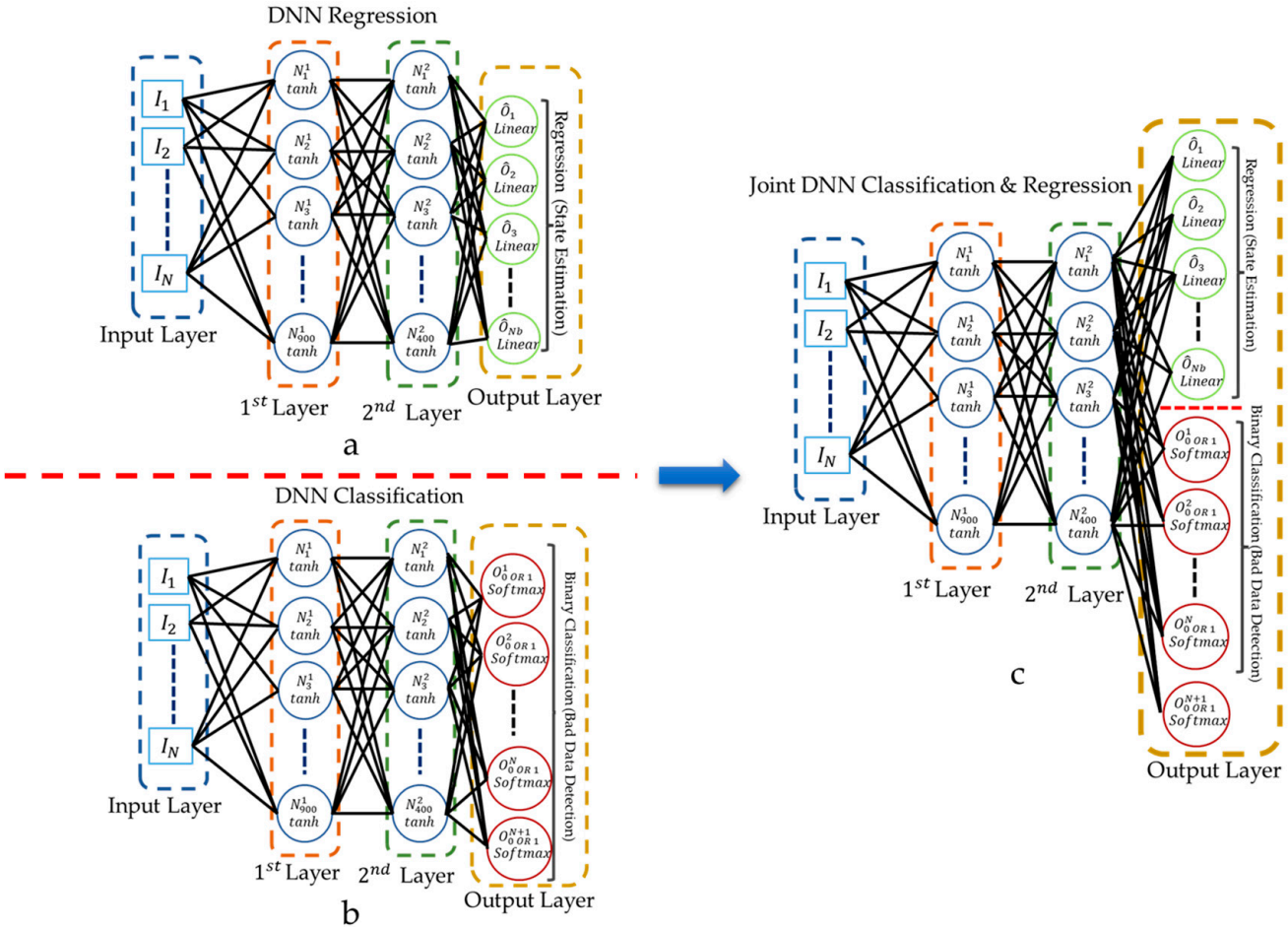

Figure 3. (**a**) A regression DNN architecture for DSSE calculation; (**b**) a binary classification DNN model for FDIA detection on measurements; (**c**) the proposed DNN model configuration to perform DSSE calculation and FDIAs on measurements simultaneously.

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), respectively:

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \cdots \\ V_N \end{bmatrix} \rightarrow V_a = \begin{bmatrix} V_{a,1} \\ V_{a,2} \\ \cdots \\ V_{a,N} \end{bmatrix} \tag{13}$$

$$I = \begin{bmatrix} I_1 \\ I_2 \\ \cdots \\ I_N \end{bmatrix} \rightarrow I_b = \begin{bmatrix} I_{b,1} \\ I_{b,2} \\ \cdots \\ I_{b,N} \end{bmatrix} \tag{14}$$

The proposed DNN model consists of input, two hidden, and output layers. The number of neurons in the input layer is equal to the number of available PMU measurements (N). The first and second hidden layers have 900 and 400 neurons, respectively. The *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 allows 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):

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

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

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

Accuracy and F_1-score, defined below, are used for binary classification outputs:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

$$F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{18}$$

In (17), TP = true positives, TN = true negatives, FP = false positives, and FN = false negatives. Precision and recall are calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{19}$$

$$\text{Recall} = \frac{TP}{TP + TN}\,. \tag{20}$$

A confusion matrix can be used to analyze 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, we highlight the entries as blue (correct) or red (incorrect).

5. Results

5.1. Case Study I

For one of the case studies, the effectiveness of the proposed method is evaluated on the IEEE 33 bus distribution network (shown in Figure 4). "True" values of electrical parameters are calculated by power flow calculations.

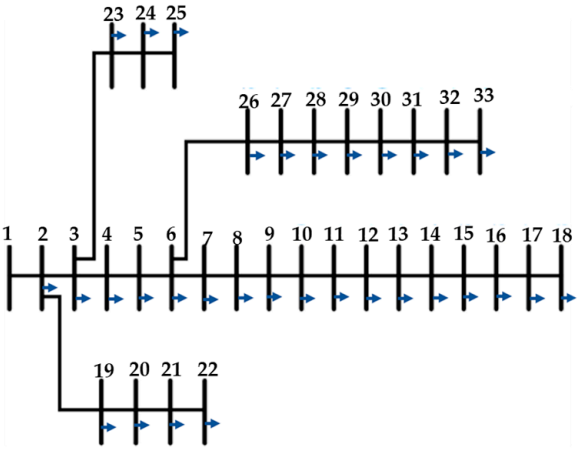

Figure 4. IEEE Standard 33 Bus Distribution System.

Measurements are randomly calculated based on their probability density function for each Monte Carlo trial. The sets of assumptions are defined as follows:

1. The number of Monte Carlo trials is chosen as $N_{MC} = 12{,}000$;
2. A Gaussian distribution, with $3\sigma = 50\%$ of the nominal value, is considered for power injection on the buses.
3. Three PMUs are assumed as the available measurement devices in the network which are located at buses 9, 16, and 31 [63]. Voltage and branch current magnitudes from each PMU are used for FDIA detection and for performing DSSE using the proposed method. For each PMU, 12,000 samples are given according to the number of Monte Carlo trials.
4. Two types of attack vectors are injected into available measurements:
   (1) $z_a = \{(1 \pm .05)z, (1 \pm 0.1)z\}$, as in [40].
   (2) $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$, where μ_z and σ_z are the average and standard deviation of each measurement vector, respectively. In this case, we assume that the attacker knows the distribution of each measurement and wants to falsify measurements based on the true measurement distribution.
5. For training and test sets, 67% and 33% of of the data were used, respectively.
6. The pseudo-measurements of active and reactive power injections and flows are generated to make the system observable and to 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: $x = [\delta_2, \cdots, \delta_N, V_1, \cdots, V_N]$, 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 $\delta_1 = 0$ and $V_1 = 1$.
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, the error of virtual measurements is considered to be 10^{-8}.

5.1.1. False Data Injection Attacks on Measurements with $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$

In this case, we inject FDIAs at random to measurement samples, and we allow 10% of each available measurement ($N_a = 1200$) to be falsified randomly by FDIAs ($z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$).

In Figure 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 we aim to identify theses erroneous measurements using the proposed method.

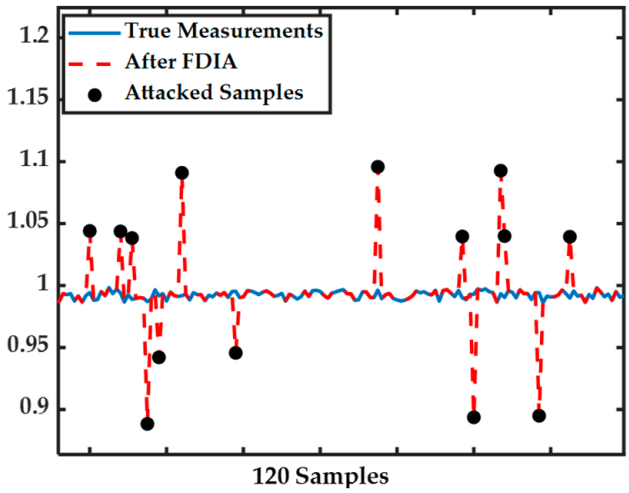

Figure 5. Original and attacked measurements for 120 samples when $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

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

Table 1 shows the results for the proposed method for the 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 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	$\lvert V_9 \rvert$	421	3	0	0	0	0	0	245	0	0	0	0	179	0
		$\lvert V_{16} \rvert$	110	260	0	0	0	0	0	0	248	0	0	0	122	0
		$\lvert V_{31} \rvert$	71	0	229	3	7	4	0	0	0	228	3	3	80	0
		$\lvert I_{9-10} \rvert$	40	0	0	266	2	2	0	0	0	0	268	0	42	0
		$\lvert I_{16-17} \rvert$	36	0	0	0	201	0	0	0	0	0	0	199	38	0
		$\lvert I_{31-32} \rvert$	1	0	0	0	0	228	0	0	0	0	0	0	229	0
	No FDIA	0		0	0	0	0	0	2076	0	0	0	0	0	0	2076
			$\lvert V_9 \rvert$	$\lvert V_{16} \rvert$	$\lvert V_{31} \rvert$	$\lvert I_{9-10} \rvert$	$\lvert I_{16-17} \rvert$	$\lvert I_{31-32} \rvert$	No FDIA	$\lvert V_9 \rvert$	$\lvert V_{16} \rvert$	$\lvert V_{31} \rvert$	$\lvert I_{9-10} \rvert$	$\lvert I_{16-17} \rvert$	$\lvert I_{31-32} \rvert$	No FDIA
		Predicted Values								Predicted Values						
		Accuracy = 0.9295 F_1-Score = 0.8885								Accuracy = 0.882 F_1-Score = 0.8142						

The results of DSSE calculation from the proposed method, the regression-only, and the WLS method are shown in Figure 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 are 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.

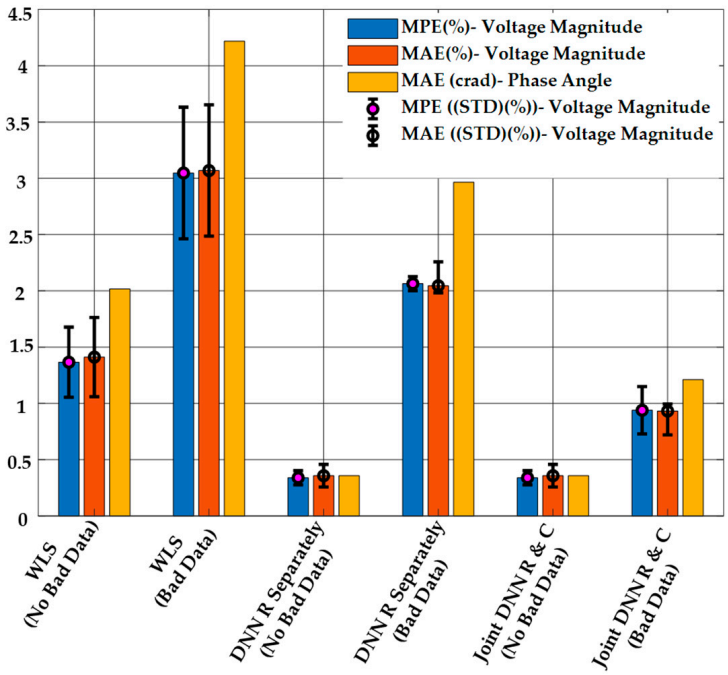

Figure 6. DSSE Results Obtained Using the Proposed Method, the Regression-only DNN Model, and WLS with or without Bad Data when $N_a = 1200$ (10%), and $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

5.1.2. False Data Injection Attacks on Measurements with $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$

In this case, we assume that the attacker knows the distribution of each measurement and wants to falsify measurements based on the true measurement distribution.

In Figure 7, 103 measurements before and after FDIAs are shown. We aim to identify falsified measurements by applying the proposed method. The FDIA detection and DSSE regression results are shown when the 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.

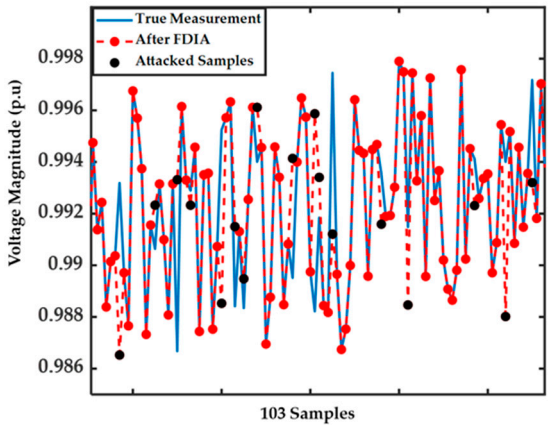

Figure 7. Original and attacked measurements for 103 samples when $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

The results in Table 2 for the proposed method and for the 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 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	$\lvert V_9 \rvert$	74	8	1	1	2	1	7	45	0	1	0	0	0	48
		$\lvert V_{16} \rvert$	6	74	0	0	0	0	3	5	44	0	0	0	0	34
		$\lvert V_{31} \rvert$	0	1	57	0	0	0	18	7	0	29	0	0	0	40
		$\lvert I_{9-10} \rvert$	2	0	0	16	2	0	57	0	0	0	1	0	0	76
		$\lvert I_{16-17} \rvert$	0	0	1	1	1	1	71	0	0	0	0	0	0	75
		$\lvert I_{31-32} \rvert$	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	3481
			$\lvert V_9 \rvert$	$\lvert V_{16} \rvert$	$\lvert V_{31} \rvert$	$\lvert I_{9-10} \rvert$	$\lvert I_{16-17} \rvert$	$\lvert I_{31-32} \rvert$	No FDIA	$\lvert V_9 \rvert$	$\lvert V_{16} \rvert$	$\lvert V_{31} \rvert$	$\lvert I_{9-10} \rvert$	$\lvert I_{16-17} \rvert$	$\lvert I_{31-32} \rvert$	No FDIA
			Predicted Values							Predicted Values						
			Accuracy = 0.9242 F_1-Score = 0.5556							Accuracy = 0.9090 F_1-Score = 0.4028						

Figure 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 are 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.

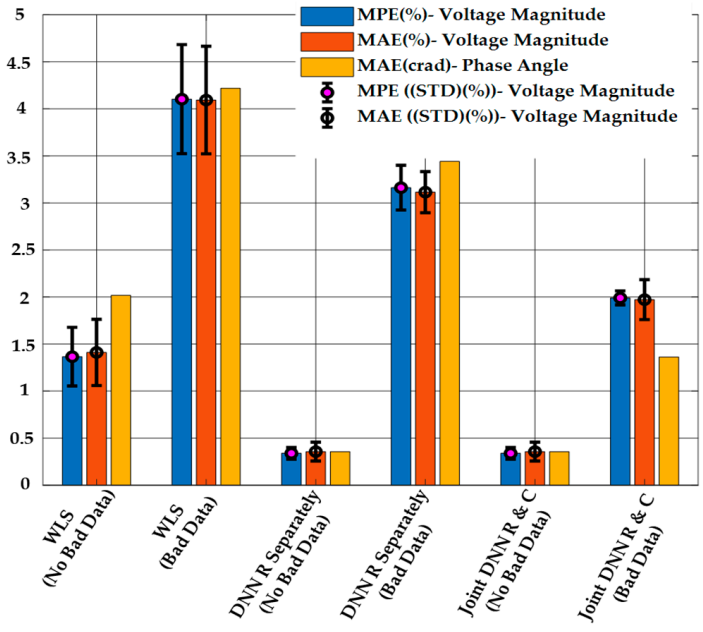

Figure 8. DSSE Results Obtained Using the Proposed Method, the Regression-only DNN Model, and WLS with or without Bad Data when $N_a = 600$ (5%) and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Table 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 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	$\lvert V_9 \rvert$	271	25	39	22	10	28	5	91	4	3	0	0	0	68
		$\lvert V_{16} \rvert$	0	312	11	5	5	3	3	13	73	1	0	0	0	60
		$\lvert V_{31} \rvert$	2	1	283	8	1	15	4	13	0	65	0	0	0	71
		$\lvert I_{9-10} \rvert$	0	0	0	244	31	22	0	4	0	0	2	0	0	143
		$\lvert I_{16-17} \rvert$	0	0	0	0	235	22	5	0	0	0	0	0	0	116
		$\lvert I_{31-32} \rvert$	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	3093
			$\lvert V_9 \rvert$	$\lvert V_{16} \rvert$	$\lvert V_{31} \rvert$	$\lvert I_{9-10} \rvert$	$\lvert I_{16-17} \rvert$	$\lvert I_{31-32} \rvert$	No FDIA	$\lvert V_9 \rvert$	$\lvert V_{16} \rvert$	$\lvert V_{31} \rvert$	$\lvert I_{9-10} \rvert$	$\lvert I_{16-17} \rvert$	$\lvert I_{31-32} \rvert$	No FDIA
			Predicted Values							Predicted Values						
			Accuracy = 0.9068 F_1-Score = 0.8557							Accuracy = 0.8393 F_1-Score = 0.4028						

Figure 9 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 MPE and MAE criteria compared to other methods in the presence of bad data. Once again, when no bad data are 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.

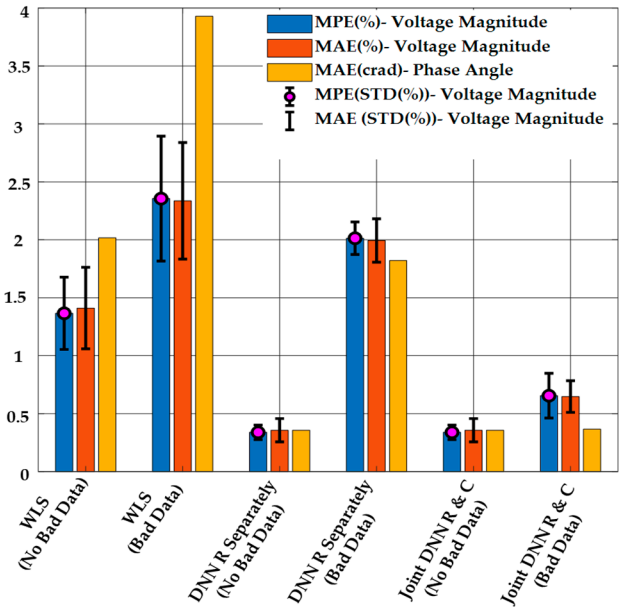

Figure 9. DSSE Results Obtained Using the Proposed Method, the Regression-only DNN Model, and WLS with or without Bad Data when $N_a = 1200$ (10%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Table 4 shows the execution time for the proposed method, regression-only method and the WLS method 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 greater than for 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. Execution time for the proposed method, regression-only method, and WLS method.

	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)	

5.2. Case Study II

The modified IEEE standard 69 bus distribution network (shown in Figure 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) [64], is utilized with the simulation time step of 1 h. Three photovoltaic (PV) panels are placed [65] on buses 21, 62, and 64 with maximum generation of 929.7 kW, 1075.2 kW, and 992.5 kW, respectively.

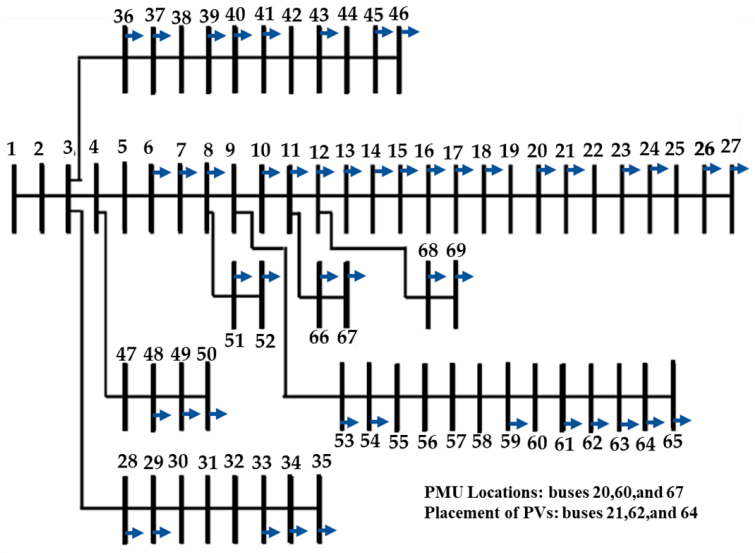

Figure 10. The modified IEEE Standard 69 Bus Distribution System.

The hourly data of power generation of a photovoltaic system are computed and adopted based on actual data for Bozeman, MT, USA [66,67]. The reactive power of bus i is defined as:

$$Q_i(t) = P_i(t) \frac{\sqrt{1 - Pf_i^2(t)}}{Pf_i(t)}, \tag{21}$$

where $Pf_i(t) \sim Unif(0.85, 0.95)$.

1. Three PMUs—located at buses 20, 60, and 67 [63]—are assumed as the available measurement devices in the network. The bus voltage and branch current magnitudes from each PMU are used for detection of FDIAs and for performing DSSE using the proposed method. We have 8760 samples with a temporal resolution of one sample per hour for each PMU, resulting in data for one year.
2. Two types of attack vector are injected into available measurements:
   (1) $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$ [40].
   (2) $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$, where μ_z and σ_z are the average and standard deviation of each measurement vector, respectively. In this case, we 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 to perform WLS calculations with the inclusion of PMU measurements.
5. Voltage magnitudes and phase angles for all buses are considered as state variable: $x = [\delta_2, \cdots, \delta_N, V_1, \cdots, V_N]$, 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 that $\delta_1 = 0$ and $V_1 = 0$.
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, the error in virtual measurements is considered to be 10^{-8}.

5.2.1. False Data Injection Attacks on Measurements with $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$

In this case, we inject the FDIAs at random to measurement samples, and we allow 10% of each available measurement ($N_a = 876$) to be falsified randomly by the FDIAs: $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

The results in Table 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 are 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 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	$\lvert V_{20}\rvert$	36	2	1	1	0	2	1	34	0	1	5	1	1	1
		$\lvert V_{60}\rvert$	1	31	3	0	0	0	2	1	31	1	0	1	1	1
		$\lvert V_{67}\rvert$	2	0	27	0	0	0	1	4	0	25	0	0	1	0
		$\lvert I_{20-21}\rvert$	0	0	0	1	0	0	31	4	0	0	3	0	0	29
		$\lvert I_{60-61}\rvert$	0	0	0	0	0	0	33	0	0	0	1	0	0	132
		$\lvert I_{67-68}\rvert$	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
			$\lvert V_{20}\rvert$	$\lvert V_{60}\rvert$	$\lvert V_{67}\rvert$	$\lvert I_{20-21}\rvert$	$\lvert I_{60-61}\rvert$	$\lvert I_{67-68}\rvert$	No FDIA	$\lvert V_{20}\rvert$	$\lvert V_{60}\rvert$	$\lvert V_{67}\rvert$	$\lvert I_{20-21}\rvert$	$\lvert I_{60-61}\rvert$	$\lvert I_{67-68}\rvert$	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 method, the regression-only method, and the WLS method are shown in Figure 11. It is clear from this figure that the proposed method has a better performance with respect to the MPE and MAE criteria compared to other methods in the presence of bad data. When no bad data are 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.

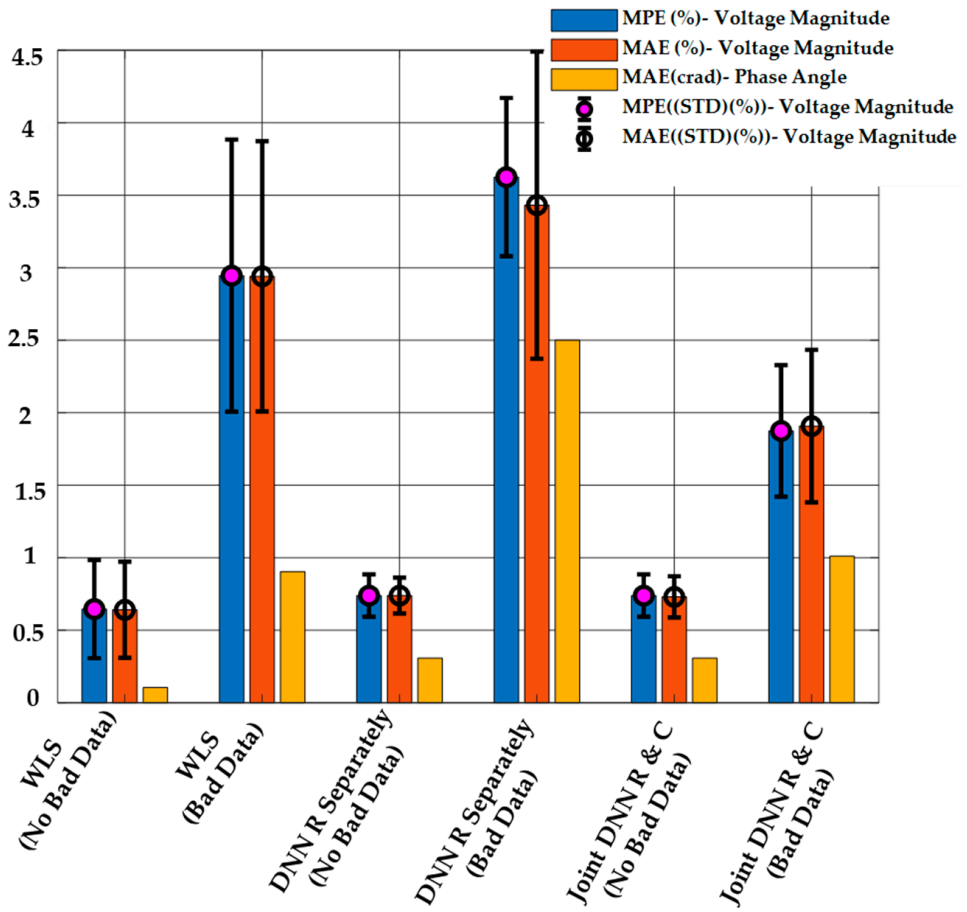

Figure 11. DSSE Results Obtained Using the Proposed Method, the Regression-only DNN Model, and WLS with or without Bad Data when $N_a = 876$, and $z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$.

5.2.2. False Data Injection Attacks on Measurements with $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$

The results in Table 6 for the proposed method and appropriate binary classification-only method show 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 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).

Figure 12 shows the results of the DSSE calculation from the proposed method, the regression-only method, and the WLS method. As it is clear from Figure 12, 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 are 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.

Table 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	$\lvert V_{20}\rvert$	22	2	2	2	0	0	4	22	3	1	2	0	0	4
		$\lvert V_{60}\rvert$	0	18	0	1	6	1	3	0	17	0	2	4	1	5
		$\lvert V_{67}\rvert$	0	0	25	1	0	4	7	1	0	24	1	0	2	9
		$\lvert I_{20-21}\rvert$	0	0	1	21	0	0	9	0	2	1	18	0	0	10
		$\lvert I_{60-61}\rvert$	0	2	0	0	18	1	10	0	2	0	0	17	2	10
		$\lvert I_{67-68}\rvert$	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
			$\lvert V_{20}\rvert$	$\lvert V_{60}\rvert$	$\lvert V_{67}\rvert$	$\lvert I_{20-21}\rvert$	$\lvert I_{60-61}\rvert$	$\lvert I_{67-68}\rvert$	No FDIA	$\lvert V_{20}\rvert$	$\lvert V_{60}\rvert$	$\lvert V_{67}\rvert$	$\lvert I_{20-21}\rvert$	$\lvert I_{60-61}\rvert$	$\lvert I_{67-68}\rvert$	No FDIA
			Predicted Values							Predicted Values						
			Accuracy = 0.9723 F_1-Score = 0.7224							Accuracy = 0.9705 F_1-Score = 0.7021						

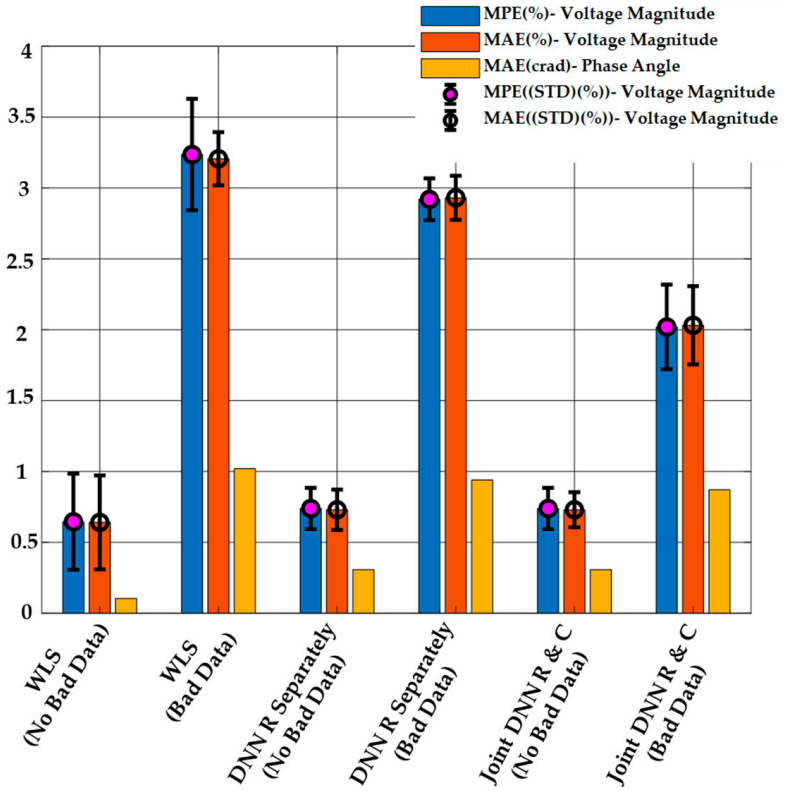

Figure 12. DSSE Results Obtained Using the Proposed Method, the Regression-only DNN Model, and WLS with or without Bad Data when $N_a = 438$ (5%), and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

Table 7 shows the results for the proposed method and the appropriate binary classification-only method, 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 *F*1-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 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 $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

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

The results of DSSE calculation from the proposed method, the regression-only method, and the WLS method are shown in Figure 13. It is clear from this figure that the proposed method has a better performance in the MPE and MAE criteria compared to the other methods in the presence of bad data. Once again, when no bad data are 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.

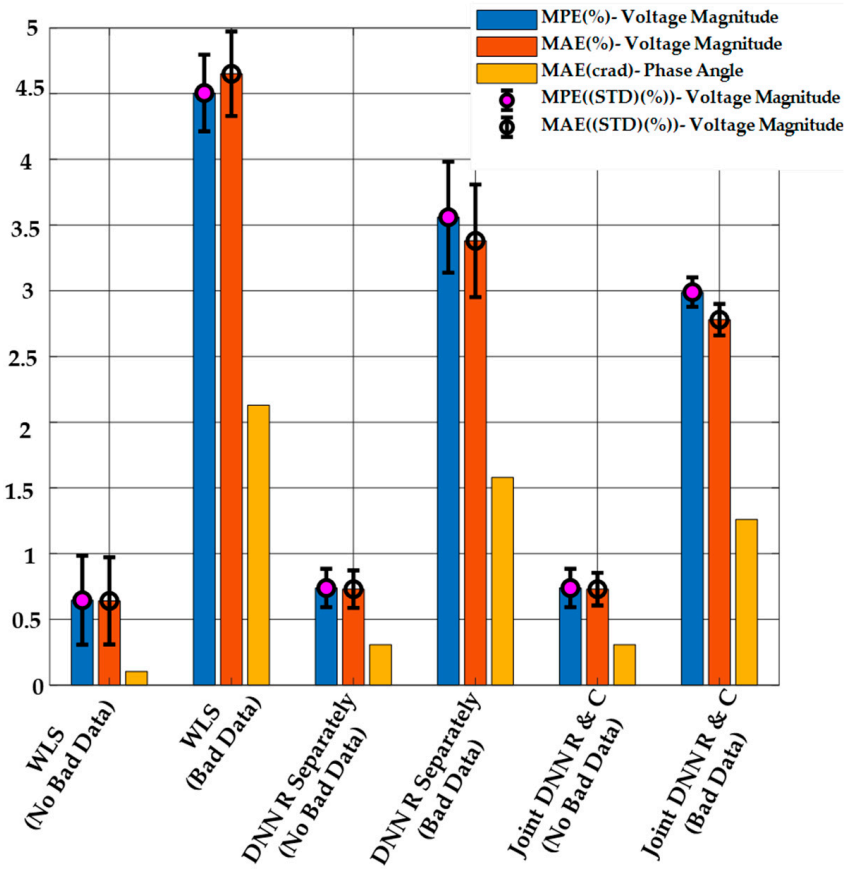

Figure 13. DSSE Results Obtained Using the Proposed Method, the Regression-only DNN Model, and WLS with or without Bad Data when $N_a = 876$, and $z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

In Table 8, the execution time for the proposed method, regression-only method, and WLS method 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. WLS execution time is ten times greater than that of the other two methods. As mentioned previously, each iteration recalculates the Jacobian matrix, which is based on physical parameters of a network, and this 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 8. Execution time for the proposed method, regression-only method, and WLS method.

	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)	

6. Conclusions

In this paper, a new method using a DNN approach is proposed to simultaneously perform DSSE calculation and FDIA detection on measurements in distribution networks. Voltage magnitudes and phase angles are defined as state vector variables in this study. 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 regression (DSSE calculation) and binary classification (FDIA detection), and the results are compared when regression and binary classification are carried out 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, we showed that DSSE calculation can be performed precisely from corrupted measurements and simultaneously identify FDIAs on corrupted measurements with high accuracy. We consider two case studies to verify the proposed method: IEEE 33-bus system without DG, and IEEE 69-bus systems with DGs. MPE and MSE values are considered to evaluate DSSE results for the proposed method, disjoint DNN method, and WLS method. Accuracy and $F1$-score are considered for evaluating the binary classification task. False data vectors are defined as being of two types: 1—$z_a = \{(1 \pm 0.05)z, (1 \pm 0.1)z\}$, and 2—$z_a = \sim \mathcal{N}(\mu_z, \sigma_z)$.

For the 33-bus case study, DSSE is performed using the proposed method, achieving 0.93% and 1.99% MPE for the first and second false data vectors, respectively. The accuracy for bad data detection is 93% and 92% for the first and second false data vectors, respectively, when 10% of each of the PMU measurements are corrupted by FDIAs. The execution time for the proposed method (min: 0.40 (s)–max: 0.63 (s)) is much faster than for the WLS method (min: 23.49 (s)–max: 30.91 (s)). For the 69-bus case study, DSSE is performed using the proposed method, achieving 1.98% and 2.01% MPE for the first and second false data vectors, respectively. The accuracy for bad data detection is 0.95 and 0.72 for the first and second false data vectors, respectively, when 10% of each of the PMU measurements are corrupted by FDIAs. The execution time for the proposed method (min: 0.15 (s)–max: 0.63 (s)) is much faster than for the WLS method (min: 43 (s)–max: 51 (s)). The difference in execution time between simultaneous and disjoint DNN models was insignificant.

Author Contributions: Conceptualization, S.R. and T.V.; methodology, S.R. and T.V.; software, S.R., T.V. and K.L.; validation, T.V., K.L. and B.M.W.; writing—original draft preparation, S.R.; writing—review and editing, T.V., B.M.W. and H.N.; visualization, S.R.; supervision, B.M.W.; project administration, H.N.; funding acquisition, H.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the US National Science Foundation under Award 1806184 and by Montana State University.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Y.; Ning, P.; Reiter, M.K. False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur.* **2011**, *14*, 21–32. [CrossRef]
2. Singh, S.K.; Khanna, K.; Bose, R.; Panigrahi, B.K.; Joshi, A. Joint-Transformation-Based Detection of False Data Injection Attacks in Smart Grid. *IEEE Trans. Ind. Inform.* **2018**, *14*, 89–97. [CrossRef]
3. Deng, R.; Xiao, G.; Lu, R.; Member, S. Defending Against False Data Injection Attacks on Power System State Estimation. *IEEE Trans. Ind. Inform.* **2017**, *13*, 198–207. [CrossRef]
4. Saldaña-González, A.E.; Sumper, A.; Aragüés-Peñalba, M.; Smolnikar, M. Advanced distribution measurement technologies and data applications for smart grids: A review. *Energies* **2020**, *13*, 3730. [CrossRef]
5. Chakhchoukh, Y.; Ishii, H. Coordinated Cyber-Attacks on the Measurement Function in Hybrid State Estimation. *IEEE Trans. Power Syst.* **2015**, *30*, 2487–2497. [CrossRef]
6. Zhuang, P.; Member, S.; Deng, R.; Liang, H. Estimation in Multiphase and Unbalanced Smart Distribution Systems. *IEEE Trans. Smart Grid* **2019**, *10*, 6000–6013. [CrossRef]
7. Xie, L.; Mo, Y.; Sinopoli, B. False Data Injection Attacks in Electricity Markets. In Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications, Gaithersburg, MD, USA, 4–6 October 2010; pp. 226–231. [CrossRef]
8. Jia, L.; Thomas, R.J.; Tong, L. Impacts of malicious data on real-time price of electricity market operations. In Proceedings of the 2012 45th Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2012; pp. 1907–1914. [CrossRef]
9. Mansouri, S.A.; Nematbakhsh, E.; Jordehi, A.R.; Tostado-Veliz, M.; Jurado, F.; Leonowicz, Z. A Risk-Based Bi-Level Bidding System to Manage Day-Ahead Electricity Market and Scheduling of Interconnected Microgrids in the presence of Smart Homes. In Proceedings of the 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I & CPS Europe), Prague, Czech Republic, 28 June–1 July 2022. [CrossRef]
10. Chen, J.; Liang, G.; Cai, Z.; Hu, C.; Xu, Y.; Luo, F.; Zhao, J. Impact analysis of false data injection attacks on power system static security assessment. *J. Mod. Power Syst. Clean Energy* **2016**, *4*, 496–505. [CrossRef]
11. Monticelli, A.; Garcia, A. Reliable Bad Data Processing for Real-Time State Estimation. *IEEE Power Eng. Rev.* **1983**, *3*, 31–32. [CrossRef]
12. Hug, G.; Giampapa, J.A. Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. *IEEE Trans. Smart Grid* **2012**, *3*, 1362–1370. [CrossRef]
13. Deng, R.; Xiao, G.; Lu, R.; Liang, H.; Vasilakos, A.V. False data injection on state estimation in power systems-attacks, impacts, and defense: A survey. *IEEE Trans. Ind. Inform.* **2017**, *13*, 411–423. [CrossRef]
14. Liu, X.; Li, Z.; Liu, X.; Li, Z. Masking Transmission Line Outages via False Data Injection Attacks. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 1592–1602. [CrossRef]
15. Ghiasi, M.; Niknam, T.; Wang, Z.; Mehrandezh, M.; Dehghani, M.; Ghadimi, N. A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future. *Electr. Power Syst. Res.* **2023**, *215*, 108975. [CrossRef]
16. Chowdhury, S.; Chowdhury, S.; Crossley, P. *Microgrids and Active Distribution Networks*; Iet Renewable Energy Series 6; The Institution of Engineering and Technology: Stevenage, UK, 2009; ISBN 9781849190145.
17. Mobtahej, M.; Esapour, K.; Tajalli, S.Z.; Mohammadi, M. Effective demand response and GANs for optimal constraint unit commitment in solar-tidal based microgrids. *IET Renew. Power Gener.* **2022**, *16*, 3485–3495. [CrossRef]
18. Wang, Z.; Zhang, B.; Mobtahej, M.; Baziar, A.; Khan, B. Advanced Reactive Power Compensation of Wind Power Plant Using PMU Data. *IEEE Access* **2021**, *9*, 67006–67014. [CrossRef]
19. Zhao, J.; Gómez-Expósito, A.; Netto, M.; Mili, L.; Abur, A.; Terzija, V.; Kamwa, I.; Pal, B.; Singh, A.K.; Qi, J.; et al. Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. *IEEE Trans. Power Syst.* **2019**, *34*, 3188–3198. [CrossRef]
20. Yang, Q.; Yang, J.; Yu, W.; An, D.; Zhang, N.; Zhao, W. On false data-injection attacks against power system state estimation: Modeling and countermeasures. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 717–729. [CrossRef]
21. Abdelmalak, M.; Venkataramanan, V.; MacWan, R. A Survey of Cyber-Physical Power System Modeling Methods for Future Energy Systems. *IEEE Access* **2022**, *10*, 99875–99896. [CrossRef]

22. Barja-Martinez, S.; Aragüés-Peñalba, M.; Munné-Collado, Í.; Lloret-Gallego, P.; Bullich-Massagué, E.; Villafafila-Robles, R. Artificial intelligence techniques for enabling Big Data services in distribution networks: A review. *Renew. Sustain. Energy Rev.* **2021**, *150*, 111459. [CrossRef]
23. Zamzam, A.S.; Fu, X.; Sidiropoulos, N.D. Data-driven learning-based optimization for distribution system state estimation. *arXiv* **2018**, *34*, 4796–4805. [CrossRef]
24. Rudin, C.; Waltz, D.; Anderson, R.; Boulanger, A.; Salleb-Aouissi, A.; Chow, M.; Dutta, H.; Gross, P.; Huang, B.; Ierome, S.; et al. Machine learning for the New York City power grid. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 328–345. [CrossRef]
25. Abdelmalak, M.; Member, S.; Hooshyar, H.; Member, S. Real-Time EMT-Phasor Co-Simulation Modeling for Large-Scale Power Grids: Challenges and Solutions. In Proceedings of the 2022 IEEE Power & Energy Society General Meeting (PESGM), Denver, CO, USA, 17–21 July 2022; pp. 1–5.
26. Anderson, B.R.N.; Ieee, M.; Boulanger, A.; Powell, W.B.; Scott, W. Adaptive Stochastic Control for the Smart Grid. *Proc. IEEE* **2011**, *99*, 14–21. [CrossRef]
27. Ozay, M.; Esnaola, I.; Yarman Vural, F.T.; Kulkarni, S.R.; Poor, H.V. Machine Learning Methods for Attack Detection in the Smart Grid. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 1773–1786. [CrossRef] [PubMed]
28. Sayghe, A.; Hu, Y.; Zografopoulos, I.; Liu, X.R.; Dutta, R.G.; Jin, Y.; Konstantinou, C. Survey of machine learning methods for detecting false data injection attacks in power systems. *IET Smart Grid* **2020**, *3*, 581–595. [CrossRef]
29. Musleh, A.S.; Chen, G.; Dong, Z.Y. A Survey on the Detection Algorithms for False Data Injection Attacks in Smart Grids. *IEEE Trans. Smart Grid* **2020**, *11*, 2218–2234. [CrossRef]
30. Esmalifalak, M.; Member, S.; Liu, L.; Member, S. Detecting Stealthy False Data Injection Using Machine Learning in Smart Grid. *IEEE Syst. J.* **2014**, *11*, 1–9.
31. Yan, J.; Tang, B.; He, H. Detection of false data attacks in smart grid with supervised learning. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 1395–1402. [CrossRef]
32. Mansouri, S.A.; Rezaee Jordehi, A.; Marzband, M.; Tostado-Véliz, M.; Jurado, F.; Aguado, J.A. An IoT-enabled hierarchical decentralized framework for multi-energy microgrids market management in the presence of smart prosumers using a deep learning-based forecaster. *Appl. Energy* **2023**, *333*, 120560. [CrossRef]
33. Foroutan, S.A.; Salmasi, F.R. Detection of false data injection attacks against state estimation in smart grids based on a mixture Gaussian distribution learning method. *IET Cyber-Phys. Syst. Theory Appl.* **2017**, *2*, 161–171. [CrossRef]
34. Joachims, T. Making Large-Scale Support Vector Machine Learning Practical. In *Advances in Kernel Methods*; MIT Press: Cambridge, MA, USA, 2022. [CrossRef]
35. Mohammadpourfard, M.; Sami, A.; Seifi, A. A statistical unsupervised method against false data injection attacks: A visualization-based approach. *Expert Syst. Appl.* **2017**, *84*, 242–261. [CrossRef]
36. Mohammadpourfard, M.; Sami, A.; Weng, Y. Identification of False Data Injection Attacks With Considering the Impact of Wind Generation and Topology Reconfigurations. *IEEE Trans. Sustain. Energy* **2018**, *9*, 1349–1364. [CrossRef]
37. Sivasangari, A.; Jyotsna, J.; Pravalika, K. SQL Injection Attack Detection using Machine Learning Algorithm. In Proceedings of the 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 3–5 June 2021; pp. 1166–1169. [CrossRef]
38. Nur, S.; Najwa, F.; Mustaffa, B. Detection of False Data Injection Attack using Machine Learning approach. *Mesop. J. Cyber Secur.* **2022**, *2022*, 38–46. [CrossRef]
39. Ramirez, I.S.; Pedro, F.; García, M. False alarm detection in wind turbine by classification models. *Adv. Eng. Softw.* **2023**, *177*, 103409. [CrossRef]
40. Mohammadpourfard, M.; Weng, Y.; Tajdinian, M. Benchmark of machine learning algorithms on capturing future distribution network anomalies. *IET Gener. Transm. Distrib.* **2019**, *13*, 1441–1455. [CrossRef]
41. Santos, R.Z.S.; Orillaza, J.R.C. Distribution system state estimator using SCADA and μ PMU measurements: An FDI attack vulnerability analysis. In Proceedings of the 2018 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Kota Kinabalu, Malaysia, 7–10 October 2018; pp. 469–474. [CrossRef]
42. Ozay, M.; Esnaola, I.; Yarman Vural, F.T.; Kulkarni, S.R.; Vincent Poor, H. Distributed models for sparse attack construction and state vector estimation in the smart grid. In Proceedings of the 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm), Tainan, Taiwan, 5–8 November 2012; pp. 306–311. [CrossRef]
43. Wang, Y.; Xing, A.; Qu, Z.; Han, X.; Dong, H.; Georgievitch, P.M. False data injection attack detection based on interval affine state estimation. *Electr. Power Syst. Res.* **2022**, *210*, 108100. [CrossRef]
44. Ganjkhani, M.; Gilanifar, M.; Giraldo, J.; Parvania, M. Integrated Cyber and Physical Anomaly Location and Classification in Power Distribution Systems. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7040–7049. [CrossRef]
45. Kamal, M.; Farajollahi, M.; Nazaripouya, H.; Mohsenian-Rad, H. Cyberattacks against Event-Based Analysis in Micro-PMUs: Attack Models and Counter Measures. *IEEE Trans. Smart Grid* **2021**, *12*, 1577–1588. [CrossRef]
46. Azimian, B.; Biswas, R.S.; Moshtagh, S.; Pal, A.; Tong, L.; Dasarathy, G. State and Topology Estimation for Unobservable Distribution Systems Using Deep Neural Networks. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 9003514. [CrossRef]
47. Abur, A.; Expósito, A.G. *Power System State Estimation: Theory and Implementation*; CRC Press: Boca Raton, FL, USA, 2004.
48. Dehghanpour, K.; Wang, Z.; Wang, J.; Yuan, Y.; Bu, F. A survey on state estimation techniques and challenges in smart distribution systems. *arXiv* **2018**, *10*, 2312–2322. [CrossRef]

49. Ahmad, F.; Rasool, A.; Ozsoy, E.; Sekar, R.; Sabanovic, A.; Elitaş, M. Distribution system state estimation-A step towards smart grid. *Renew. Sustain. Energy Rev.* **2018**, *81*, 2659–2671. [CrossRef]
50. Liao, H.; Milanović, J.V. Pathway to cost-efficient state estimation of future distribution networks. In Proceedings of the 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, USA, 17–21 July 2016. [CrossRef]
51. Džafić, I.; Gilles, M.; Jabr, R.A.; Pal, B.C.; Henselmeyer, S. Real time estimation of loads in radial and unsymmetrical three-phase distribution networks. *IEEE Trans. Power Syst.* **2013**, *28*, 4839–4848. [CrossRef]
52. Gao, Y.; Yu, N. State estimation for unbalanced electric power distribution systems using AMI data. In Proceedings of the 2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 23–26 April 2017. [CrossRef]
53. Muscas, C.; Sulis, S.; Angioni, A.; Ponci, F.; Monti, A. Impact of different uncertainty sources on a three-phase state estimator for distribution networks. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 2200–2209. [CrossRef]
54. Majumdar, A.; Pal, B.C. A three-phase state estimation in unbalanced distribution networks with switch modelling. In Proceedings of the 2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI), Kolkata, India, 8–10 January 2016; pp. 474–478. [CrossRef]
55. Kong, X.; Chen, Y.; Yong, C.; Ma, X.; Kong, J. Stepwise robust distribution system state estimation considering PMU measurement. *J. Renew. Sustain. Energy* **2019**, *11*, 025506. [CrossRef]
56. Huang, Y.F.; Werner, S.; Huang, J.; Kashyap, N.; Gupta, V. State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid. *IEEE Signal Process. Mag.* **2012**, *29*, 33–43. [CrossRef]
57. Simeone, O. An Introduction to Quantum Machine Learning for Engineers. *Found. Trends Signal Process.* **2022**, *16*, 1–223. [CrossRef]
58. Neranjan Thilakarathne, N.; Mohan, K.K.; Surekha, L.; Hussain, A. Smart Grid: A Survey of Architectural Elements, Machine Learning and Deep Learning Applications and Future Directions. *J. Intell. Syst. Internet Things* **2021**, *3*, 32–42. [CrossRef]
59. Zhang, D.; Han, X.; Deng, C. Review on the research and practice of deep learning and reinforcement learning in smart grids. *CSEE J. Power Energy Syst.* **2018**, *4*, 362–370. [CrossRef]
60. B, P.; Pham, Q.-V.; Liyanage, M.; Deepa, N.; VVSS, M.; Reddy, S.; Maddikunta, P.K.R.; Khare, N.; Gadekallu, T.R.; Hwang, W.-J. Deep Learning for Intelligent Demand Response and Smart Grids: A Comprehensive Survey. *arXiv* **2021**, 1–25.
61. Moradzadeh, A.; Mohammadpourfard, M.; Konstantinou, C.; Genc, I.; Kim, T.; Mohammadi-Ivatloo, B. Electric Load Forecasting under False Data Injection Attacks using Deep Learning. *Energy Rep.* **2022**, *8*, 9933–9945. [CrossRef]
62. Radhoush, S.; Vannoy, T.; Whitaker, B.M.; Nehrir, H. Simultaneous State Estimation and Bad Data Detection on PMU Measurements in Active Distribution Power Networks. In Proceedings of the 2022 North American Power Symposium (NAPS), Salt Lake City, UT, USA, 9–11 October 2022; pp. 1–6. [CrossRef]
63. Lee, K.Y.; Park, J.S.; Kim, Y.S. Optimal placement of pmu to enhance supervised learning-based pseudo-measurement modelling accuracy in distribution network. *Energies* **2021**, *14*, 7767. [CrossRef]
64. Office of Energy Efficiency & Renewable Energy (EERE). Commercial and Residential Hourly Load Profiles. Available online: https://openei.org/datasets/files/961/pub/ (accessed on 23 February 2023).
65. Sedighizadeh, M.; Jahangir, M.; Gandomkar, M.; Esfandeh, S. Distributed generation location and capacity effect on voltage stability of distribution network. *Int. Conf. Math. Methods Comput. Tech. Electr. Eng. Proc.* **2010**, *25*, 89–94.
66. Joseph, A. *Shaw ORSL Weather Station*; The Montana State University Weather Station: Bozeman, MT, USA; Available online: https://www.montana.edu/orsl/weather.html (accessed on 23 February 2023).
67. NREL. PVWatts Version 5 Manual. Available online: https://pvwatts.nrel.gov/ (accessed on 23 February 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.