energies

MDPI

Article

State Estimation Fusion for Linear Microgrids over an Unreliable Network

Mohammad Soleymannejad [1,†], Danial Sadrian Zadeh [1,†], Behzad Moshiri [1,2,*], Ebrahim Navid Sadjadi [3], Jesús García Herrero [3] and Jose Manuel Molina López [3]

[1] School of Electrical and Computer Engineering, University of Tehran, Tehran 1417614411, Iran; soleymannejad@ut.ac.ir (M.S.); daniyal.sadriyan@ut.ac.ir (D.S.Z.)
[2] Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[3] Department of Informatics, Universidad Carlos III de Madrid, 28903 Madrid, Spain; 100367078@alumnos.uc3m.es (E.N.S.); jgherrer@inf.uc3m.es (J.G.H.); molina@ia.uc3m.es (J.M.M.L.)
* Correspondence: moshiri@ut.ac.ir
† These authors contributed equally to this work.

Abstract: Microgrids should be continuously monitored in order to maintain suitable voltages over time. Microgrids are mainly monitored remotely, and their measurement data transmitted through lossy communication networks are vulnerable to cyberattacks and packet loss. The current study leverages the idea of data fusion to address this problem. Hence, this paper investigates the effects of estimation fusion using various machine-learning (ML) regression methods as data fusion methods by aggregating the distributed Kalman filter (KF)-based state estimates of a linear smart microgrid in order to achieve more accurate and reliable state estimates. This unreliability in measurements is because they are received through a lossy communication network that incorporates packet loss and cyberattacks. In addition to ML regression methods, multi-layer perceptron (MLP) and dependent ordered weighted averaging (DOWA) operators are also employed for further comparisons. The results of simulation on the IEEE 4-bus model validate the effectiveness of the employed ML regression methods through the RMSE, MAE and R-squared indices under the condition of missing and manipulated measurements. In general, the results obtained by the Random Forest regression method were more accurate than those of other methods.

Keywords: cyberattack; data fusion; estimation fusion; internet of things; Kalman filter; machine learning; packet loss; smart microgrid; state estimation

Citation: Soleymannejad, M.; Sadrian Zadeh, D.; Moshiri, B.; Sadjadi, E.N.; García Herrero, J.; Molina López, J.M. State Estimation Fusion for Linear Microgrids over an Unreliable Network. *Energies* **2022**, *15*, 2288. https://doi.org/10.3390/en15062288

Academic Editor: Ali Mehrizi-Sani

Received: 30 January 2022
Accepted: 16 March 2022
Published: 21 March 2022

1. Introduction

Microgrids, including distributed energy resources (DERs), are currently being seriously considered due to their greenhouse gas emissions being considerably lower than power plants. Microgrids are therefore considered suitable for connection to the main grid and result in reduced transmission costs and losses [1]. In contrast to the benefits that they can offer, their natural patterns of power generation cause critical challenges for the operational stability of the power system [2].

With the proliferation of machine-learning (ML) methods, a new window has opened on the scalability and accuracy of smart microgrids for attack generation, detection, and mitigation strategies [3]. The consideration of unreliability in the communication network is motivated by several real-world inevitable factors that can be categorized into two major aspects. The first one is the inherent lossy behavior of the channel, such as missing packets, link failures of the internet infrastructure, packet delays, device failures, disturbances, and attenuation of the sent and received signals [4,5]. The second one is the external entities that are mainly considered as threats to the integrity of transferred data. These entities are mostly observed to manipulate the sensed information [6].

In this regard, cyberattacks have always caused major economic, social, and technical challenges, such as blackouts, the manipulation of smart sensor readings, and the alteration of the estimated profiles of load in power systems [7].

Regarding the sources of unreliability listed above, which adversely affect the integrity and, more importantly, the accuracy of transmitted data, the communication network between the microgrid and the energy management system (EMS) needs to resort to unreliability-mitigation approaches [8]. It is, therefore, crucial to estimate the microgrid states obtained from the received information and, consequently, apply these mitigation approaches for monitoring enhancement [9].

1.1. Related Works

Various tools and approaches have been adopted to address the problems of smart grids with communication systems regarding state estimation. For instance, state estimation based on Kalman filter (KF) over wireless sensor networks (WSNs) with fading channels was proposed in [10]. Generally, this type of state estimation using centralized techniques requires a massive amount of computation and communication resources. It is primarily prone to be threatened by central point failures that may result in disastrous blackouts.

Furthermore, a distributed approach for network-based sensor fusion using KF was presented in [11] in order to address the communication problems. In this approach, the fusion center combines the local estimators employing a set of premeditated weighting factors. A similar approach with the help of ordered weighted averaging (OWA) operator is considered in [12] for missing measurements. Besides, in order to calculate a proper weighting factor, a recursive algorithm based on the weighted density function was developed in [13] for reliable communication channels.

Approaches inspired by the covariance intersection for fusion via aggregation weights were used in [14–16]. Rana [17] examined an attack-resistant algorithm in which the attack is automatically ignored and the state estimation process continues. This acts as a grid eye that monitors the entire power system. Since the altered sensor measurements are the primary sources of uncertainty, the novelty of this algorithm lies in modifying the correction step of the KF with the help of the residual saturation function, the calculation of which depends upon the weighting factor and the residual dynamics.

In addition, to study the injecting power stabilization of nonlinear DC microgrids with constant power demands, a third-degree cubature Kalman filter (CKF) was suggested in [18] to reduce the influence of the noisy measurement and the noisy network on the system's information. For higher-order DC microgrids with a large number of sources and constant power loads (CPLs), the proposed CKF method is resilient against system uncertainty and noisy surroundings and has a short computing time. Likewise, a similar study was performed in [19] for microgrid state estimation under cyberattack using spherical simplex radial cubature Kalman filter (SSRCKF).

Moreover, the authors of [20] proposed the use of a satellite system based on the internet of things (IoT) to transmit measurement information from sensing devices in the grid to the EMS. This proposal stems from the need to accurately and continuously monitor the power systems to ensure a reliable service for customers, mainly consisting of a suitable voltage and frequency. In their proposed structure, the IoT elements, such as sensors and actuators, are utilized to collect microgrid data that are relatively enormous in scale.

Most of the state estimation methods tend to utilize weighted least squares (WLS) in order to address state estimation issues under cyberattack conditions [21–23]. A sequential injected false data detection in smart grids was presented in [24]. This method implements a generalized likelihood-based ratio centralized detector with a cumulative sum algorithm.

Moslemi [25] offered a fast, decentralized technique for cyberattack detection in smart grids on the basis of maximum likelihood (ML) estimation that makes use of power grids' near-chordal sparsity to create an efficient framework for solving the ML estimation problem through a modified Newton method. By preventing data exchange across areas, the

suggested decentralization ensures the privacy of the utilities and decreases the complexity of ML problems by downsizing.

Concerning the data manipulation threats, Mustafa [26] discussed a robust control system for distributed frequency and voltage regulation of AC microgrids. An attack detection technique utilizing a Kullback–Leibler (KL) divergence-based criterion is provided for each DER to identify any misbehavior on its adjacent DERs. The computed KL divergence factors are then used to produce belief values that indicate the validity of the received information, which is then recommended as an attack mitigation strategy.

The authors of [27] investigated leveraging the strength of the KF to overcome the problem of false data injection (FDI) into the sensed information. Likewise, intentional injections of false synchrophasor measurements can lead to erroneous control actions, thus, jeopardizing the security and reliability of transmission networks. Hence, the authors of [28] presented a multisensor track-level fusion-based model prediction (TFMP) as a solution to this problem. Each monitoring node utilizes a Kalman-like particle filter (KLPF)-based smoother to recover the initial correlation information regarding attacked oscillation parameters. The KLPF-based smoother is divided into subsystems to decrease its computational load. As a result, the initial oscillatory state estimations are improved.

Rana [29] proposed a distributed state estimation algorithm that considers the packet loss in the smart grid environment. The novelty of this algorithm lies in the method of calculating the best weighting factor for estimating the global states. At the same time, the same author [30] proposed a novel consensus filter-based dynamic state estimation algorithm for distributed state estimation in a modern power system. The algorithm relies on the mean squared error (MSE) and semidefinite programming methods to calculate the optimal local gain between the actual states and the estimated ones. The consensus gain is also determined with the help of a convex optimization process with a given sub-optimal local gain.

The authors of [31] investigated the application of decentralized state estimation to the combined heat and power system (CHPS) regarding possible communication failures among different energy systems. The proposed approach, namely the relaxed alternating direction method of multipliers (R-ADMM), provides efficiency in computational costs. In addition, Qu [32] studied the problem of estimating the dynamic states of islanded microgrids affected by the fading measurement. The proposed approach is a recursive state estimation scheme whose estimation error is assured to be within a certain upper limit, allowing for the online monitoring of islanded microgrids.

The exploitation of a delay-universal-based error correction coding scheme is discussed in [33] for achieving reliable and real-time state estimates from an IoT-based unstable microgrid by alleviating the error impacts of the communication channel. Rana [34] considered the problem of robust estimation of the network states and proposed a technique for estimating the state of a power distribution system containing multiple synchronous generators in the presence of a network attack. The proposed method is based on graph theory and optimal filter, which improves the performance of the network state estimation process.

Wang [35] studied the possibility of applying deep learning for estimating the states of power systems and proposed a physically-guided deep learning (PGDL) method inspired by autoencoders and deep neural networks (DNNs) regarding their ability in learning temporal correlations. The proposed PGDL is known to be data- and physically-driven. In addition, Tanvir [36] discussed the estimation and control of a DC microgrid based on wind power by employing an adaptive KF to estimate the rotor flux and an artificial neural network (ANN) to estimate the rotor velocity. Consequently, the robustness of parameter uncertainty due to adaptive mechanisms is improved.

The authors of [37] provided state estimation of a real-time power system based on data using a deep-ensemble-learning algorithm. The proposed setup consists of several parallel ResNetD stacked as base-learners and multivariate-linear regression as a meta-learner. This setup uses historical measurements and states for training based on power

system states, voltage amplitude, and phase estimates. The trained model is then made use of to predict the states of the power system in real-time using real-time measurements.

The authors of [38] used a conditional deep belief network (CDBN) to distinguish the behavior aspects of FDI attacks with historical measurement data and use the recorded features to detect FDI attacks in real-time. Accordingly, the suggested detection technique effectively relaxes assumptions about possible attack situations, thus, achieving high accuracy.

Furthermore, Deng [39] explored the vulnerabilities of smart distribution systems to FDI attacks by first presenting a local state-based linear distribution system state estimation (DSSE) for multiphase and imbalanced distribution systems from the attacker's perspective. The probabilities of a successful FDI attack are also calculated mathematically. The case study results demonstrate the possibility for attackers to launch FDI attacks in a practical multiphase and imbalanced smart distribution system with varying levels of effort.

1.2. Primary Contributions

The related works discussed above have common characteristics. They only considered one of the communication imparities, such as packet dropouts, cyberattacks, or FDI. With a new perspective for using machine-learning (ML) regression methods, as estimation fusion methods, for KF-based state estimators, the key contributions of this paper are as follows:

1. The current study considers three sources of uncertainty simultaneously applied to the measurement data. These sources are packet dropout, IoT channel noise, and cyberattack.
2. The current study applies different ML regression methods, employed as data fusion methods, to fuse the KF-based state estimates and compare the results for higher accuracy. The first reason is to determine if these easy and fast methods can substitute for methods requiring high computational costs and complex mathematical formalization. The second reason is to determine whether aggregating state estimates using these methods can contribute to a more accurate estimate of state than accomplished by each KF alone.
3. This study also illustrates the inefficacy of the weighted averaging operators, which are used in many prior studies, through fusing the state estimates by the dependent ordered weighted averaging (DOWA) operator.

1.3. Manuscript Layout

The remainder of this manuscript is arranged as follows. Section 2 presents a model of a linear microgrid, followed by the structure of a lossy IoT-based communication network and the proposed method. Section 3 is dedicated to the results of the simulation while Section 4 provides discussion and analysis. Lastly, the paper is concluded in Section 5, followed by a statement of future work.

2. Materials and Methods

This section introduces the model under consideration, the properties of the communication network and its associated issues, as well as the proposed approaches to solve the issues.

2.1. Linear Microgrid Model

In this study, as shown in Figure 1, we assumed that $N = 4$ DERs are connected to the IEEE 4-bus test feeder at the points of common coupling (PCCs) [40,41]. Each DER requires its voltage to be controlled in order to maintain its reference values. The dynamics of the system are linear and are given as follows:

$$\dot{x}(t) = Ax(t) + Bu(t) + \omega(t), \tag{1}$$

where the state vector $x(t) = [v_1 \;\; v_2 \;\; v_3 \;\; v_4]^T$ is the PCC voltage vector, A is the state transition matrix, B is the input matrix, $u(t)$ is the control input, and $\omega(t)$ is the process noise given by

$$\omega \sim \mathcal{N}(0, Q), \tag{2}$$

in which Q is the covariance matrix of the process noise.

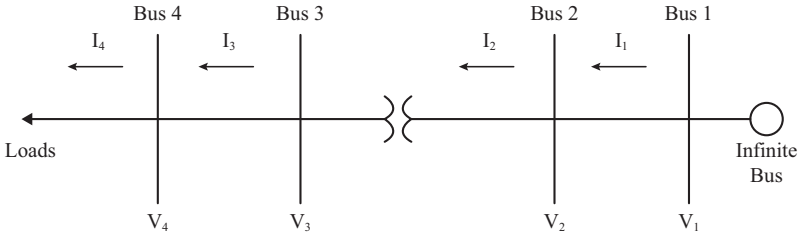

Figure 1. Illustration of the IEEE 4-bus model.

In order to work with this model, discretization is required. Hence, the employed rule of discretization is as shown below.

$$\dot{x}(t) = \frac{x(t + \Delta t) - x(t)}{\Delta t} = Ax(t) + Bu(t), \tag{3}$$

$$x(t + \Delta t) = \big(Ax(t) + Bu(t)\big)\Delta t + x(t). \tag{4}$$

By considering that the sampling process occurs every Δt seconds, (4) can be rewritten as follows:

$$x(k+1) = (I + A\Delta t)x(k) + (B\Delta t)u(k), \tag{5}$$

where [41]

$$A = \begin{bmatrix} 175.9 & 176.8 & 511 & 103.6 \\ -350 & 0 & 0 & 0 \\ -544.2 & -474.8 & -408.1 & -828.8 \\ -119.7 & -554.6 & -968.8 & -1077.5 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.8 & 334.2 & 525.1 & -103.6 \\ -350 & 0 & 0 & 0 \\ -69.3 & -66.1 & -420.1 & -828.8 \\ -434.9 & -414.2 & -108.7 & -1077.5 \end{bmatrix}.$$

Hence, the final representation is as follows:

$$x_{k+1} = A_d x_k + B_d u_k + \omega_k. \tag{6}$$

Now that the dynamics of power system states have been determined, the smart sensors that can directly sense the system states need to be considered, thus, forming an observation or measurement model. In this study, four observation stations, which measure all four voltages, were considered to reduce uncertainties in the sensed data by creating redundancy. Hence, the measurement equation can be written as follows:

$$y_k^i = Hx_k + v_k^i, \tag{7}$$

where y_k^i is the information vector observed by the i-th observation station at time step k, H is the system's observation matrix, and v_k^i is the measurement noise vector for the i-th observation station caused by the distributed wireless sensors [42]. Therefore, the measurement noise can be written as follows:

$$v^i \sim \mathcal{N}(0, R^i), \tag{8}$$

where R^i is the covariance matrix of the measurement noise of the i-th observation station.

2.2. Unreliable Communication Channel Structure

In this subsection, first, the occurrence of cyberattacks in terms of the information generated in a microgrid is considered. A lossy and unreliable channel is then presented for transmitting this manipulated data, and its behavior is discussed.

2.2.1. Cyberattack

Suppose that a malicious agent, also known as the attacker, attacks at time τ and deliberately tampers with the sensor readings by a time-varying value F_k. Subsequently, the measurement change can be formalized as follows [24]:

$$y_k^i = \begin{cases} Hx_k + v_k^i & k < \tau\Delta t, \quad \text{(9a)} \\ Hx_k + F_k^i + v_k^i & k \geq \tau\Delta t. \quad \text{(9b)} \end{cases}$$

As mentioned in [43], F_k^i can be decomposed into two parts, as shown below:

$$F_k^i = Hc_k^i + \beta_k^i. \tag{10}$$

The parameter β_k^i is the only informative part of the injected false data that is detectable. However, the term Hc_k^i can purportedly bypass the monitoring system's security measures as it is difficult to differentiate from Hx_k per each state estimation. Therefore, as the attackers avoid being detected, they mostly attempt to disguise attack vectors in the column space of the measurement matrix H. If attackers have seamless knowledge of the network characteristics (i.e., the values of matrix H, the network topology, line admittances, and transformer tap ratio) and can exploit any sensor they desire, then they would be able to perform stealth attacks that bypass the security system [43].

Fortunately, this is not the usual case in microgrids due to certain factors in real-world situations. Firstly, most attackers often acquire the network's characteristics by offline learning of the network entities' behavior over a long period while the power grid configuration could change over time. Thus, the attackers are not likely to collect and analyze the network information in real-time. Secondly, the attackers generally have an inadequate level of access to manipulate a large enough number of sensing devices [43].

This paper assumes that the attackers are unable to predict the system dynamics, as a consequence of which, they are unable to access the distribution parameters of the measurements. Consequently, the injection of false data has significant effects on the distribution of the measured data, which decreases the correlation between the data acquired by each sensing station [44,45]. This leads to the assumption that some components of β^is are nonzero, meaning that the trace of the attack vector always exists. Thus, the change in the sensed data can be redrafted as follows:

$$\beta_k^i = \begin{cases} b_m^k = 0 & m \notin \Omega \quad k < \tau\Delta t, \quad \text{(11a)} \\ |b_m^k| > \gamma & m \in \Omega \quad k \geq \tau\Delta t, \quad \text{(11b)} \end{cases}$$

where γ is an agreed value that describes the lower bound for the measurement tolerance that draws security attention, and the set Ω includes randomly chosen columns of the state-space vectors that are to be attacked in each time step k. Here, the objective is to abate the effect of false data injected undesirably through an attack vector β^i that is exploiting the measurements of i-th observation station soon after its occurrence at $\tau\Delta t$ [24].

2.2.2. IoT Network and Packet Loss

The sensed information from each observation station has to be sent to the EMS. To accomplish this purpose, an observation station needs to send the gathered data over a

collection of relay nodes (i.e., the internet) due to the long distance between the DERs and the EMS.

As the sensing devices are mostly assumed to be wirelessly communicating with the sensing station, there needs to be an integrated communication setup incorporated into the microgrid to facilitate the data exchange process between these two mentioned sides. Consequently, the IoT network is the concept capable of addressing this issue.

In the case of this study, a WSN powered by 5G technology is considered as the appropriate IoT-based communication network, relaying the sensed information to the estimation station. In the IoT network, by the act of a uniform quantizer, the information signals are transformed into corresponding bit sequences, B_k, in each time step k. These bit sequences are then modulated as a signal, S_k, by proceeding through the modulation process using the binary phase-shift keying (BPSK) technique. Eventually, S_k is transmitted over the internet to the EMS.

Considering the cyberattack discussed earlier in this section, the measurement model (7) can be restated as follows:

$$y^i_{st,k} = Hx_k + v^i_k + \beta^i_k, \tag{12}$$

where $y^i_{st,k}$ is the measurement vector to be transmitted from the i-th observation station under the condition of cyberattacks. As discussed earlier, each $y^i_{st,k}$ is transformed into a bit sequence B^i_k and then modulated to S^i_k. The transmission of the measurement data over an IoT network, by its intrinsic characteristics, may cause delayed or lost measurements due to packet dropouts [46]. Regarding the mentioned losses, cyberattacks, quantization, and modulation, the received measurements transmitted from the i-th observation station to the corresponding estimator in the time step k (12) can be redrafted as follows:

$$y^i_{rd,k} = \alpha^i_k S^i_k + \alpha^i_k e_k, \tag{13}$$

in which e_k is the additive white Gaussian noise (AWGN) and α^i_k is the IoT-based packet loss parameter. This parameter is modeled as shown below [4]:

$$\alpha^i_k = \begin{cases} 1 \pm \delta & \text{with probability of } \lambda_k, \quad (14a) \\ 0 & \text{with probability of } 1 - \lambda_k, \quad (14b) \end{cases}$$

where λ_k is the packet arrival rate of each observation station at the EMS at time step k, and δ is fundamentally a fraction of sent packet value that is randomly added or subtracted during the process of transmission in order to practically model the noisy nature of IoT-based communication networks. Finally, the log-maximum a posteriori (Log-MAP) decoding method is used to decode each received signal [47]. The decoded outputs are then followed by demodulation and dequantization, thus, being prepared to be fed to the state estimators.

2.3. Proposed Method

In this section, the proposed methods are discussed in two steps, namely state estimation and estimation fusion: (1) processing all received data with distributed KF estimators and (2) fusing the state estimates based on the received data for each state using various ML regression methods to obtain more accurate results. An illustration of the proposed distributed state estimation and estimation fusion regarding cyberattacks and packet loss is shown in Figure 2.

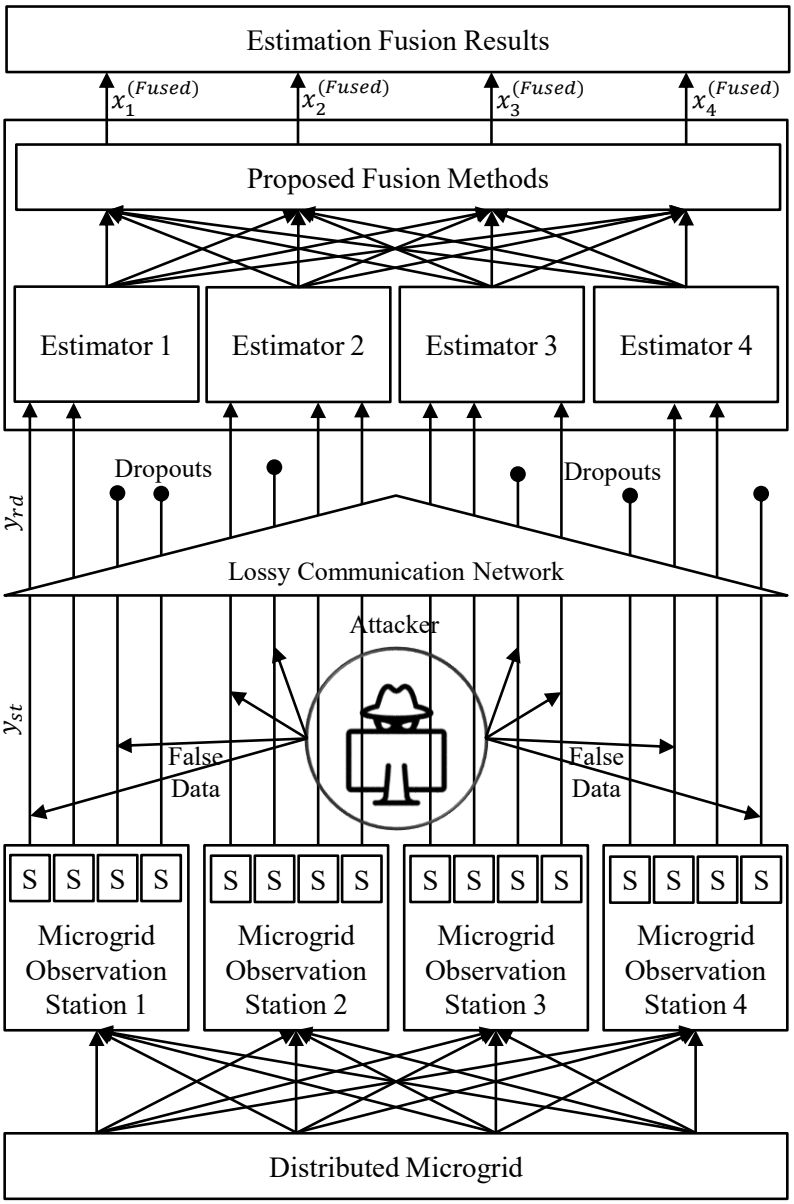

Figure 2. The proposed distributed microgrid state estimation and estimation fusion under cyberattack and packet loss (a significantly modified version of the benchmark structure given in [48]).

2.3.1. Distributed State Estimation

The KF algorithm for each estimator that is responsible for estimating the states based on the measurements received from the observation stations is written as follows [49]:

$$\hat{x}^i_{k|k-1} = A_d \hat{x}^i_{k-1|k-1} + B_d u_k, \tag{15}$$

$$P^i_{k|k-1} = A_d P^i_{k-1|k-1} A^T_d + , \tag{16}$$

$$K^i_k = P^i_{k|k-1} H^T \left(H P^i_{k|k-1} H^T + R \right)^{-1}, \tag{17}$$

$$\hat{x}^i_{k|k} = \hat{x}^i_{k|k-1} + K^i_k \left[y^i_{rd,k} - H \hat{x}^i_{k|k-1} \right], \tag{18}$$

$$P^i_{k|k} = (I - K^i_k H) P^i_{k|k-1}. \tag{19}$$

In these equations, $\hat{x}^i_{k-1|k-1}$ is the state estimate of the previous step by the i-th estimator, $P^i_{k-1|k-1}$ is the predicted covariance matrix of the previous step, $\hat{x}^i_{k|k}$ is the current state estimate, and K^i_k is the Kalman gain.

2.3.2. Estimation Fusion

Once the estimates of a state variable are obtained, they should be fused to yield a single value representing the final estimation of that state variable for each time step k. In this study, a novel perspective of using the ML regression methods is considered to fuse the state estimates and compare the methods with popular data fusion approaches, such as the dependent ordered weighted averaging (DOWA) operator as well as multi-layer perceptron (MLP).

In the literature, several ML methods and algorithms have been surveyed with regard to their applicability in the context of fusion [50–52]. Here, the ML regression methods of interest are Linear, Decision Tree, AdaBoost, XGBoost, Gradient Boosting, Random Forest and Voting. The reason for selecting these algorithms is due to their fast training and testing process, as timeliness is one of the benefits of data fusion [53].

Regression is a linear model that assumes a linear relationship between the input values and gives a single variable as the output. To mathematically define this, the output y given by a linear regression model for a set of input data $\{x_1, x_2, \ldots, x_n\}$ is calculated as follows:

$$y = \sum_{i=1}^{n} w_i x_i + b. \tag{20}$$

The task of estimation fusion is to find weights such that the fusion of state estimates via the obtained weights shows the most resemblance to the actual data—that is, the values of the fusion of estimated states are expected to have the least possible difference with the actual state values, and this is acquired by training a suitable model. In regression methods, weights are computed dynamically based on solving an optimization problem to minimize a loss function.

Combining multiple regression methods to improve the prediction accuracy can be called multi-regressor fusion (MRF). Two well-known types of MRFs are bagging and boosting methods.

Bagging is mainly a method of generating multiple variants of a predictor (e.g., the regression model in this study) and leveraging them to achieve an aggregated prediction—the aggregation averages over the multiple variants when predicting a numerical outcome. Experiments on actual and simulated datasets using regression trees have shown that bagging can considerably improve the accuracy of a model [54].

Among the most successful bagging methods in this study is the Random Forest regression method, which uses a large number of decision tree regressors bundled together. After performing the bootstrapping process on the input data and feeding the bootstrapped samples to the regression trees, this method aggregates the results to produce the final output. Assuming there are L tree regressors and, correspondingly, L subsets of bootstrapped data, the ensemble learner has to assign weights w_l $(l = 1, \ldots, L)$ to each of the aforementioned results and then operate as an aggregator as shown below:

$$S_L = \frac{1}{L} \sum_{l=1}^{L} w_l \hat{z}_l, \tag{21}$$

where $\hat{z}_l$ denotes the output of l-th tree regressor. The main goal here is to find each w_l such that S_L approaches the actual label of the sample, causing a decrease in the relative error.

On the other hand, boosting methods are defined as functional gradient descent algorithms that work iteratively to optimize a cost function over the function space by selecting a function (weak hypothesis as regressor) [55].

More specifically, the weak regressors are added one after another, exploring each iteration to find the best pair to add to the current ensemble model. In other words, the model S_L in the l-th iteration is recurrently defined. For example, in AdaBoost, the following is considered [56]:

$$s_l = s_{l-1} + c_l w_l, \tag{22}$$

where c_l and w_l are optimized in a way that s_l is the best improvement over s_{l-1} and better fits the training data. The optimization problem to find c_l and w_l can be written as follows:

$$(c_l, w_l) = \arg\min_{c,w} E(s_{l-1} + c_l w_l), \tag{23}$$

in which $E(\cdot)$ is the fitting error of the given model. Similarly, for Gradient Boosting regression method, (23) can be modified to achieve the weights and coefficients using the gradient descent approach as follows:

$$s_l = s_{l-1} - c_l \nabla_{s_{l-1}} E(s_{l-1}). \tag{24}$$

Considering the descriptions on finding appropriate weights to combine outputs in boosting algorithms, some of the most popular algorithms of this type, such as adaptive boosting (AdaBoost), Gradient Boosting and XGBoost, were selected in this study for fusion comparison.

In addition to the aforementioned methods, three other methods were employed: (1) the DOWA operator, which is introduced in [57]; (2) an MLP model, as a universal function approximator [58] to obtain a better overview of how efficient our proposed methods are; and (3) a voting regressor over three of the best aforementioned methods with the lowest error values.

The main underlying reason behind employing these methods is that most of the previous works, as discussed in the introduction, made use of the weighted averaging operators. This means that the fusion of the input set $\{x_1, x_2, \ldots, x_n\}$ is performed by n number of weights that follow the rule below:

$$\begin{cases} \sum_{i=1}^{n} w_i = 1, & \text{(25a)} \\ w_i \in [0,\ 1] \quad i = 1, \ldots, n. & \text{(25b)} \end{cases}$$

However, the problem with these methods is that the aggregated result is always between the minimum and maximum values in the input set $\{x_1, x_2, \ldots, x_n\}$. If the true value is out of this range, these methods have no chance of increasing the estimation accuracy. OWA operators are in this category, and this paper employs the DOWA operator to justify the reason for the rejection of the weighted averaging operators in such applications. Nevertheless, these methods have other applications that lead to excellent results (the interested reader is referred to [59,60]).

3. Results

To construct a framework that simulates the entire process as closely as possible to real-world conditions, this work implements a framework encompassing state generation, measurement sensing, lossy IoT-based communication network with cyberattack and packet loss, state estimation, and, finally, estimation fusion, as illustrated in Figure 2.

In this experiment, 150 time steps (snapshots) of measurements for each of the four states from each observation station are generated; these measurements experience cyberattacks and go through a lossy IoT communication network. Having been received by the fusion center, the measurements are then fed to the KFs for the purpose of state estimation. In other words, four KFs estimate each of the four states of the microgrid. Hence, a total amount of 16 state estimates in every time step is generated.

This operation is repeated 100 times to perform a Monte Carlo (MC) simulation, resulting in a training dataset consisting of 15,000 samples of estimated and true states—that is, a matrix with 20 columns and 15,000 rows. Furthermore, 25 runs of MC simulations are also performed to create a test dataset. The parameters for these performed processes are mentioned in Table 1.

Table 1. Table of the parameters used for the simulations.

Parameter	Description	Value
H	Observation matrix	$I_{4\times4}$
Q	Process noise cov.	$1 \times 10^{-7} \times I_{4\times4}$
R^1	Measurement noise cov. for station 1	$1 \times 10^{-6} \times I_{4\times4}$
R^2	Measurement noise cov. for station 2	$2 \times 10^{-6} \times I_{4\times4}$
R^3	Measurement noise cov. for station 3	$3 \times 10^{-6} \times I_{4\times4}$
R^4	Measurement noise cov. for station 4	$4 \times 10^{-6} \times I_{4\times4}$
λ	Package arrival rate	0.95
γ	Lower bound for false data	2
δ	Fraction of sent packet value	0.05
Δt	Sampling time	0.0001 s

Three measures are employed in order to evaluate the accuracy of each estimation; these include MAE, RMSE and R-squared. According to Table 2, which provides estimation errors of the KFs (before fusion) on the test dataset, both MAE and RMSE measures for states 2 and 4 have their lowest values for the first estimator; however, for states 1 and 3, both measures show a more reliable performance for the second estimator. This difference is due to the stochasticity and uncertainty imposed on the measurements by the communication network.

Table 2. Kalman filter-based estimation errors on the test dataset.

State	Station	MAE	RMSE	R^2
$\hat{x}_1$	1	3.2963	5.0993	0.9720
	2	3.0632	4.9420	0.9737
	3	4.7249	6.7034	0.9517
	4	4.5799	6.0224	0.9610
$\hat{x}_2$	1	2.7704	5.9068	0.9860
	2	6.5434	13.4811	0.9271
	3	5.8461	13.7188	0.9245
	4	5.2894	12.8139	0.9341
$\hat{x}_3$	1	2.0455	2.4221	0.9854
	2	1.6878	2.2850	0.9870
	3	1.8052	2.2576	0.9873
	4	2.5473	3.1690	0.9750
$\hat{x}_4$	1	1.3977	2.6451	0.9953
	2	5.5890	12.9268	0.8878
	3	5.4177	12.4675	0.8956
	4	5.0777	12.1761	0.9004

After the estimation phase, data must be fused by using the methods discussed previously. In the fusion phase, the regression models will be trained on the previously created training dataset through five-fold cross-validation process and tested on the test dataset for fusion evaluation. The error values, computed by the three aforementioned measures, are then reported for the fusion results in order to provide a comparison between the methods of interest.

Regarding the training process of the fusion methods of interest, one model of each regression method per state number is trained using the training dataset. Their parameters are tuned empirically, using a random grid search with five-fold cross-validation. This implementation was performed using the scikit-learn library in the Python programming language. For the MLP model, four MLPs, one for each state, with four layers were trained with four input nodes and one output node, iterating for 120 epochs.

The training features of each MLP are the ReLU activation function for the first three layers, Linear activation function for the output layer, Adam optimizer, and MSE loss function. The MinMaxScaler is also used to preprocess the data before feeding the data to the MLPs. The training process is performed using the TensorFlow library in Python. In addition, the DOWA operator, which does not need training, is leveraged to show the inefficacy of the weighted averaging operators in such harsh situations.

The expectation is to achieve MAE and RMSE values that are less than those of the KF estimators. It is also expected to obtain higher values for R-squared, which determines how well the results fit the target data (i.e., the ground truth). Note that, in order to see if the regression weights are reproducible, the training process was run several times for each method. The weights were initialized based on the normal distribution each time. Thus, the initial weights differ each time the process of training is performed, and the same results were obtained each time. Figures 3–9 illustrate the training and validation loss function values for all states for all the trainable methods.

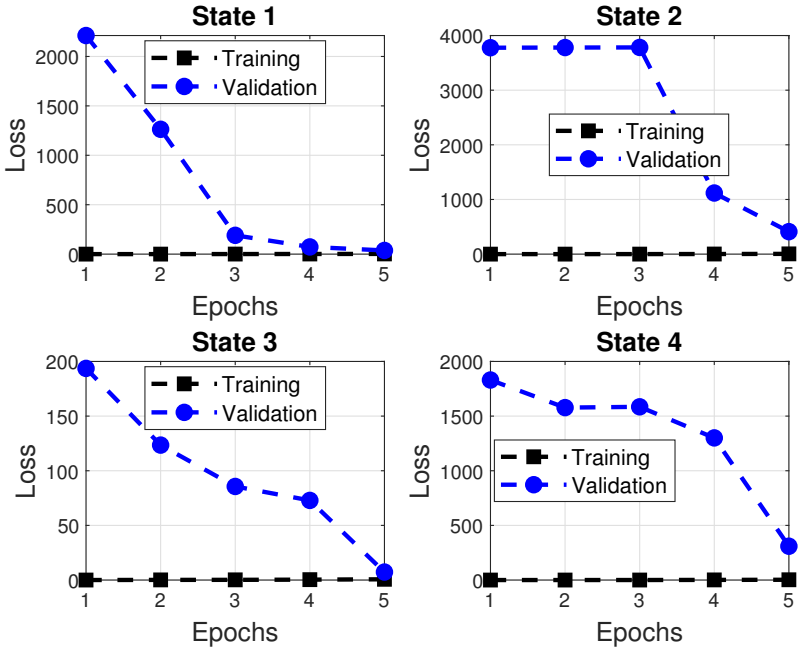

Figure 3. Training and validation MSE losses for AdaBoost.

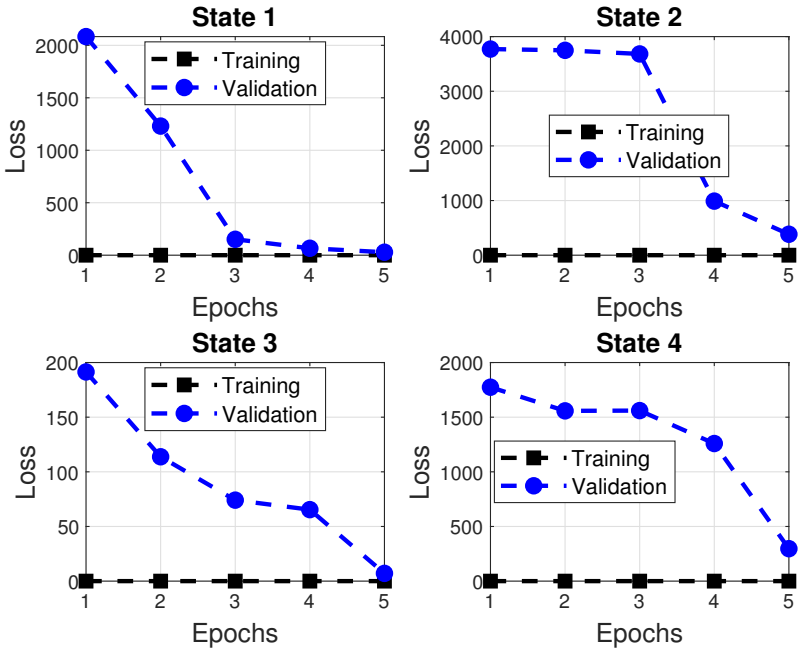

Figure 4. Training and validation MSE losses for Decision Tree.

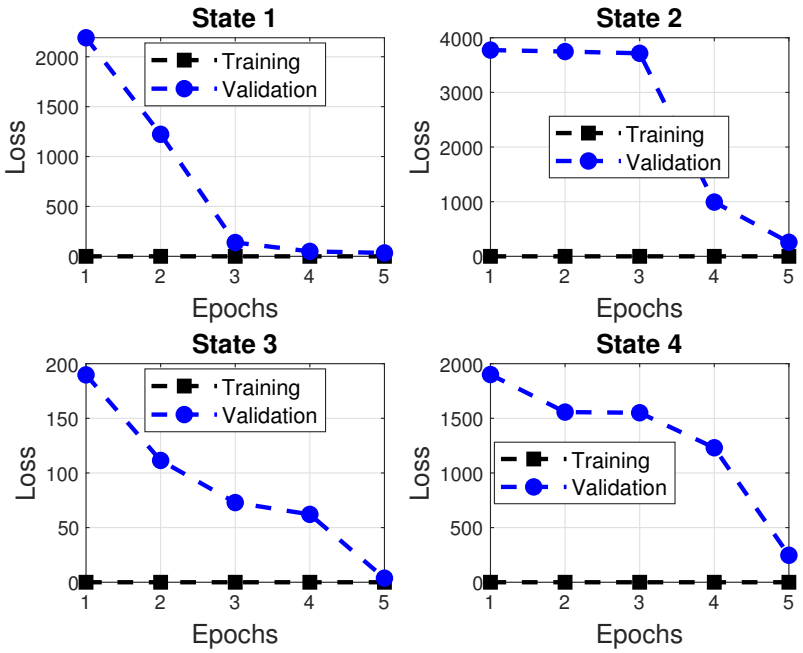

Figure 5. Training and validation MSE losses for Gradient Boosting.

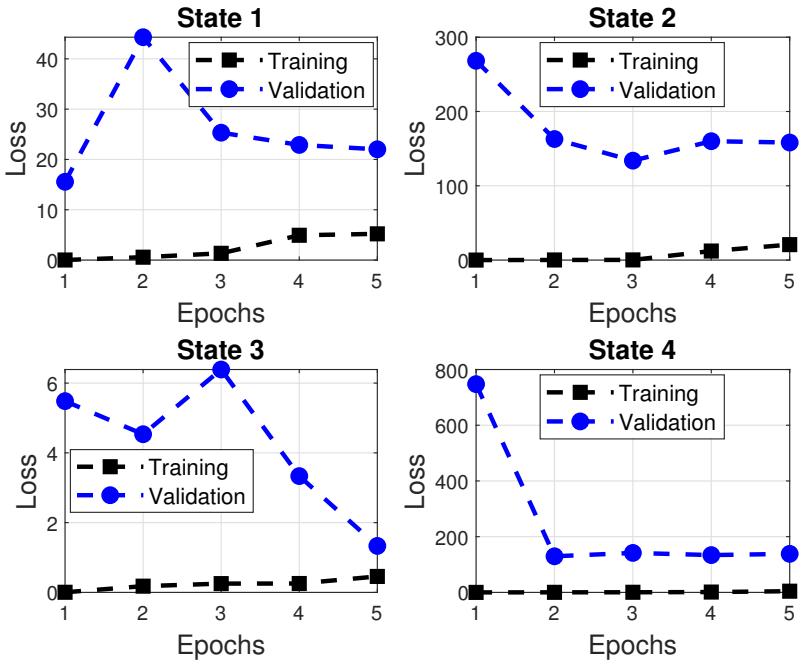

Figure 6. Training and validation MSE losses for Linear Regression.

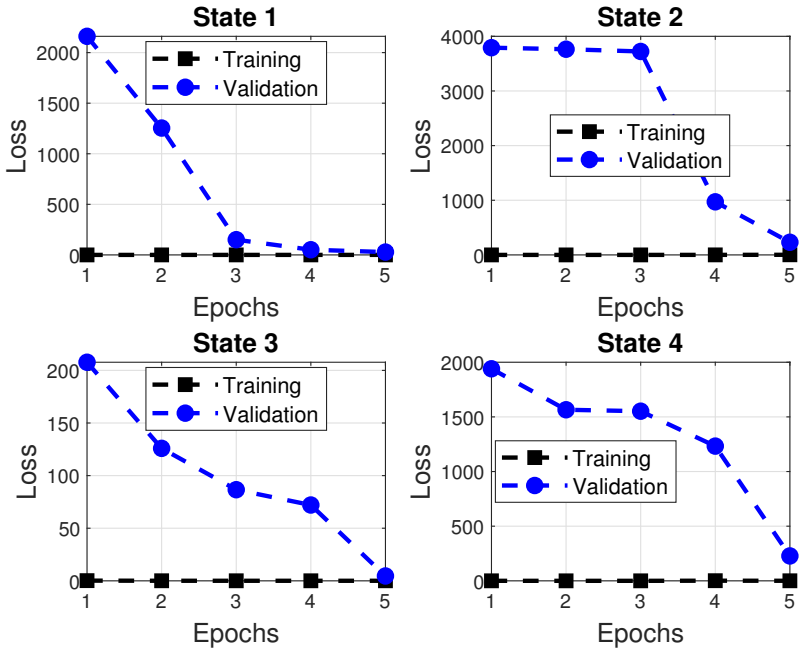

Figure 7. Training and validation MSE losses for Random Forest.

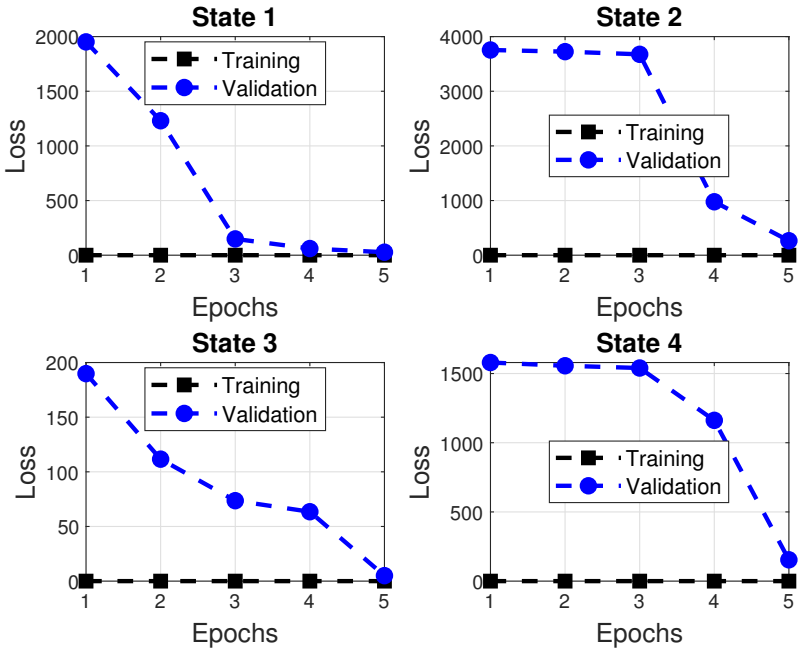

Figure 8. Training and validation MSE losses for XGBoost.

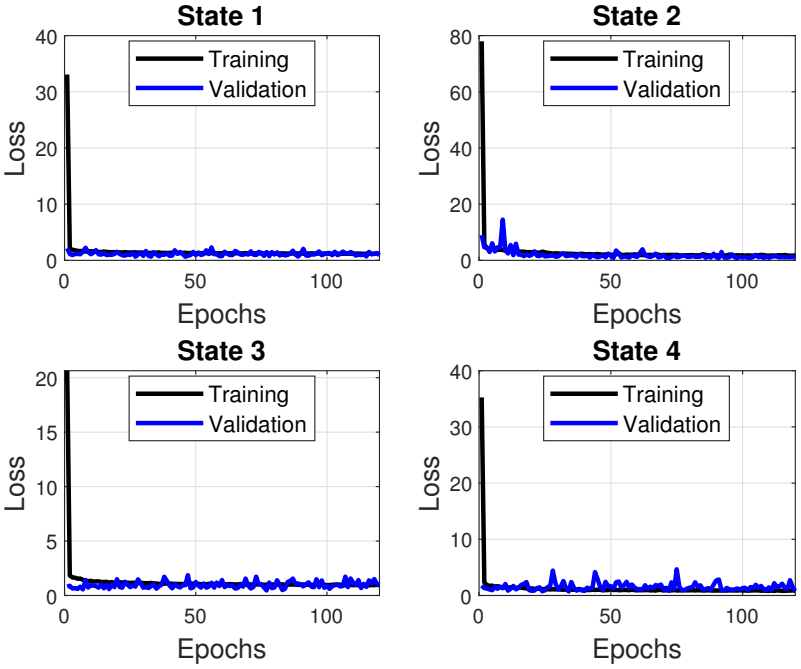

Figure 9. Training and validation MSE losses for MLP.

According to the evaluation measures utilized, the majority of the employed methods, with the exception of the DOWA operator, surpassed the findings for the estimation phase, as shown in Table 3. For example, for the state 1, the best error obtained by estimation was 3.0632 in terms of MAE and 4.942 in terms of RMSE. The worst results obtained by the regression methods of interest for this state were 2.087 and 2.59, respectively, which were from AdaBoost. This pattern can be observed for other states as well.

Table 3. Estimation fusion errors on the test dataset.

Method	$\hat{x}_1^{(Fused)}$			$\hat{x}_2^{(Fused)}$			$\hat{x}_3^{(Fused)}$			$\hat{x}_4^{(Fused)}$		
	MAE	**RMSE**	$\mathbf{R^2}$	**MAE**	**RMSE**	$\mathbf{R^2}$	**MAE**	**RMSE**	$\mathbf{R^2}$	**MAE**	**RMSE**	$\mathbf{R^2}$
MLP	0.8730	1.1630	0.9978	0.7817	1.1147	0.9989	0.7746	1.1175	0.9931	0.7242	1.0401	0.9985
Voting Reg.	0.9212	1.2468	0.9973	0.8192	1.1869	0.9988	0.7908	1.0903	0.9938	0.7342	1.0016	0.9985
Random Forest Reg.	0.9296	1.2987	0.9970	0.7909	1.1711	0.9989	0.7746	1.0845	0.9937	0.7000	0.9684	0.9985
Gradient Boosting Reg.	1.0026	1.3354	0.9969	1.0634	1.4952	0.9980	0.8487	1.1633	0.9933	0.9179	1.2327	0.9979
XGBoost Reg.	0.9904	1.3297	0.9971	0.8873	1.2941	0.9987	0.8674	1.1981	0.9928	0.7746	1.0852	0.9982
AdaBoost Reg.	2.0871	2.5905	0.9879	3.3988	4.2964	0.9827	1.4784	1.9343	0.9825	2.5024	3.0333	0.9805
Linear Reg.	1.4753	2.1712	0.9923	2.2026	3.7697	0.9900	0.9733	1.3927	0.9929	1.8457	2.7007	0.9919
Decision Tree Reg.	1.1926	1.7118	0.9953	1.1342	1.6855	0.9977	1.0405	1.5540	0.9889	0.9334	1.3450	0.9972
DOWA Operator	13.2437	15.7244	0.6106	19.3938	23.2389	0.2535	9.1564	11.1311	0.5228	14.3409	17.9817	0.3774

4. Discussion

One of the most important outcomes of this experiment is that the bagging method, i.e., the Random Forest method, had the best results for all states compared to the other methods, surprisingly, in close competition with the MLP. It even outperformed the MLP in RMSE and R-squared measures for state 3 and all measures for state 4. This performance of the Random Forest is presumed to be due to its outstanding ability for interpolation since the test data points rest among the training data points.

As a bagging method, Random Forest has another significant advantage over the boosting methods, specifically AdaBoost. It is known to be less afflicted with the noise of data, which leads to a better generalization by reducing variance, and this is mainly due to the fact that the error of generalization reaches a limit regardless of the increase in the number of decision trees.

AdaBoost, on the other hand, is expected to show weak performance when the dataset is highly noisy, which was the case for this study. This is because it spends too much time learning the most extreme cases and distorting the results, followed by worsened error rates when computational complexity is a concern. Accordingly, the results achieved by Random Forest were much better than for AdaBoost, which is shown in Table 3.

Among the boosting methods, Gradient Boosting and XGBoost demonstrated an acceptable performance competing for better error values for different states. The most referenced advantage of the XGBoost method is its speed compared to other boosting methods. In addition, as expected, the Decision Tree method almost ranks after these discussed methods since it is mainly known to be beneficial in situations where certain feature values are missing or the noise is relatively high.

The Linear Regression method is also considered prior to AdaBoost in this study. The main reason for its observed deficiencies in certain cases could be because of a possible high rate of partial non-linearity in the relationship between the data points and the target values. A particular case is where the increasing or decreasing trends of received measurement values reach a turning point, and voltages tend to change toward the opposite direction as shown in Figures 10–11.

Packet loss is more observable and devastating on the received data in these particular areas and thus more difficult to alleviate. In the cases where the trend is linear, Linear Regression operates normally, outperforming AdaBoost for most states and the DOWA operator method for all states regarding all error measures. This is because of the discussed limitations of the DOWA operator. After ensuring that the results of Random Forest and the two mentioned boosting algorithms are sufficient, the Voting regression method is built

using these three methods. The results, as reported, are not superior to the Random Forest method but better than the other two for all states.

Finally, both the estimation and fusion phases overcame the high noise variance of the received measurements even before the cyberattack started to interfer at $k = 15$. This means that, even if the attacker tends to alter the value of β, the employed methods of interest can overcome this effect. No value for β would be able to stay undetected and still cause a severe problem or noticeable trace in the estimation and fusion results. This is because the attacker can only use values of β of a specific range to avoid being detected. In addition, these methods relax the destructive influence of packet dropouts, which results in zero-valued received measurements in certain time steps. The fusion results for one MC simulation for each state are shown in Figures 10–17.

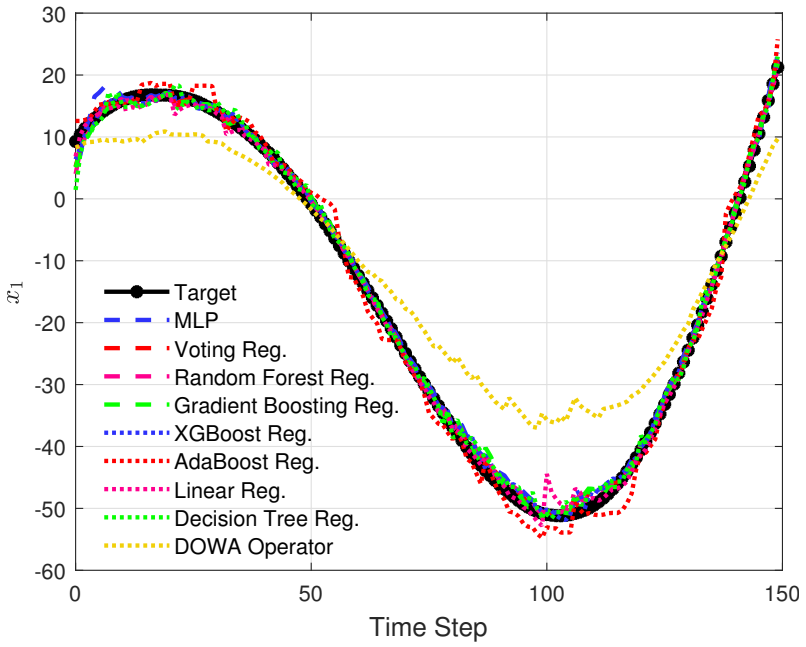

Figure 10. Fusion results for one Monte Carlo simulation for state 1.

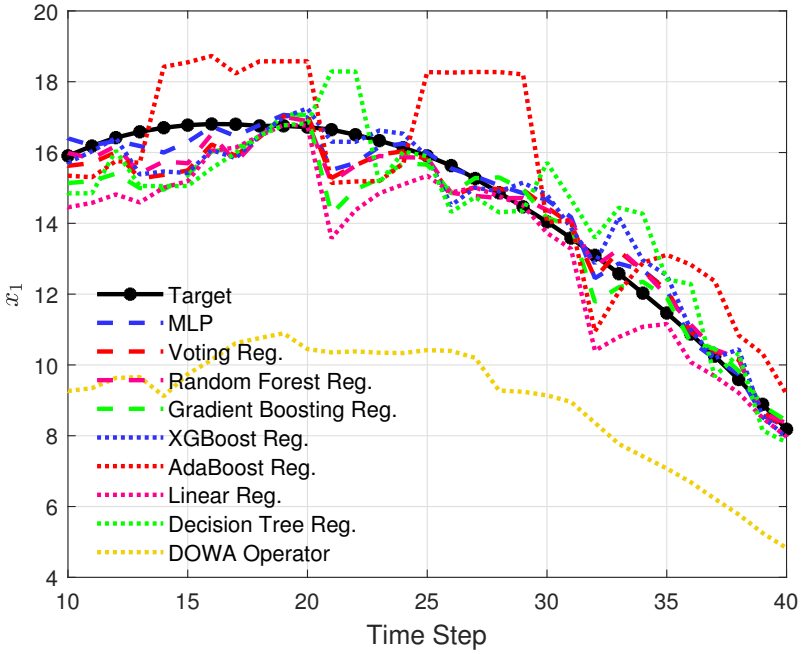

Figure 11. Fusion results for one Monte Carlo simulation for state 1 (zoomed in).

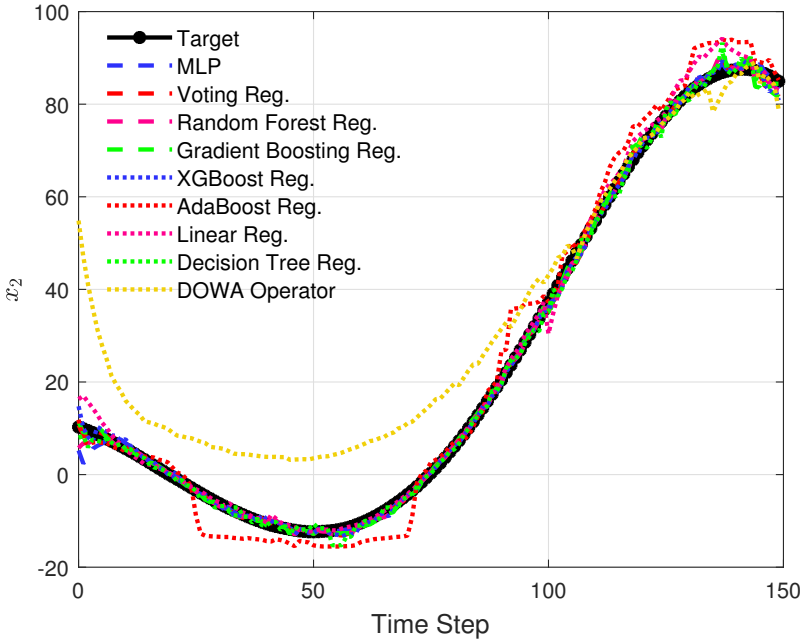

Figure 12. Fusion results for one Monte Carlo simulation for state 2.

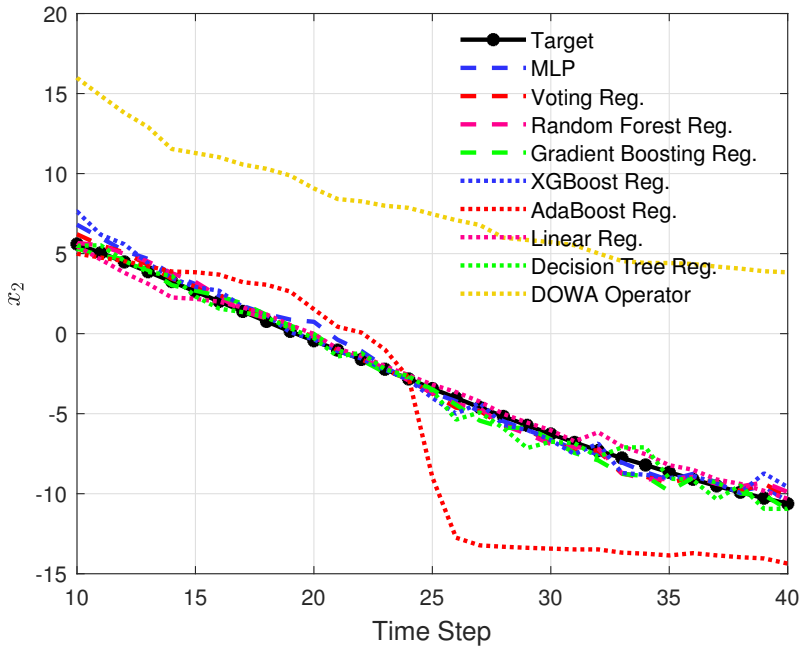

Figure 13. Fusion results for one Monte Carlo simulation for state 2 (zoomed in).

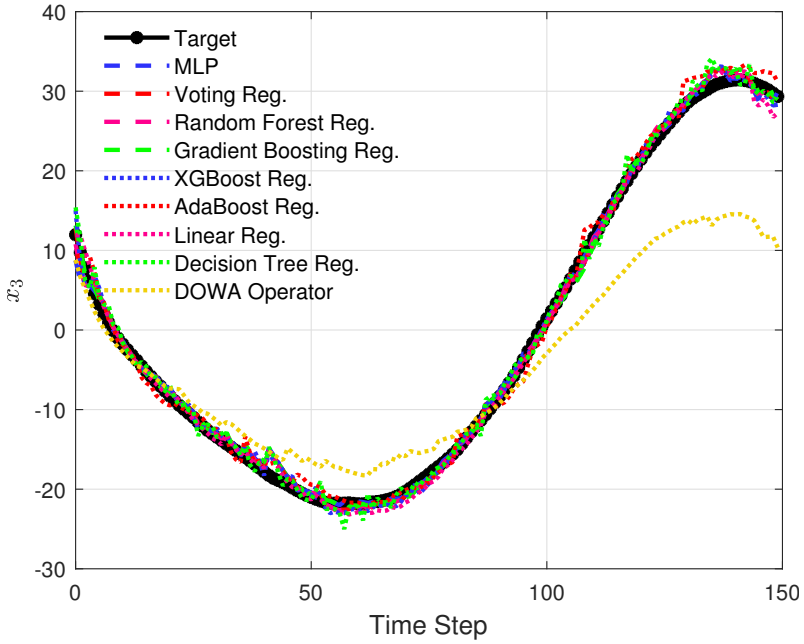

Figure 14. Fusion results for one Monte Carlo simulation for state 3.

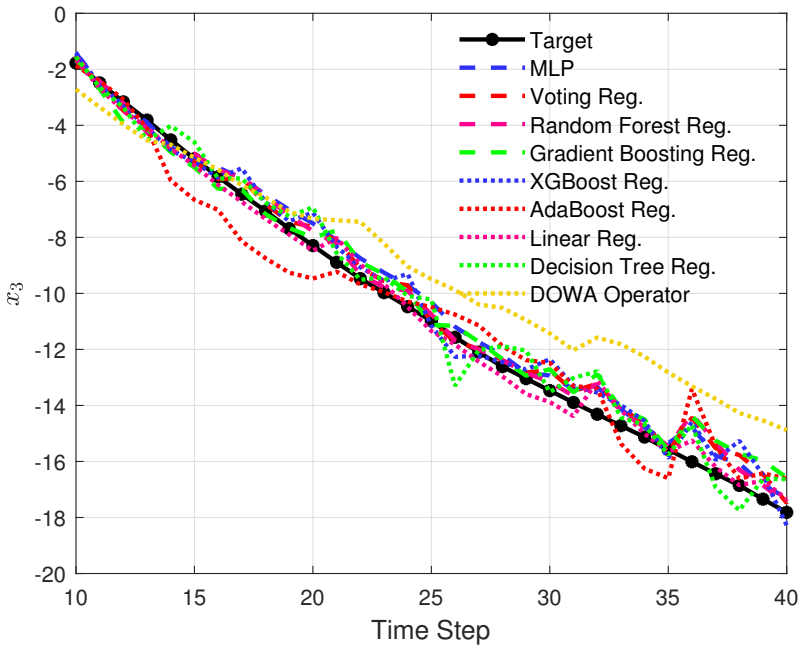

Figure 15. Fusion results for one Monte Carlo simulation for state 3 (zoomed in).

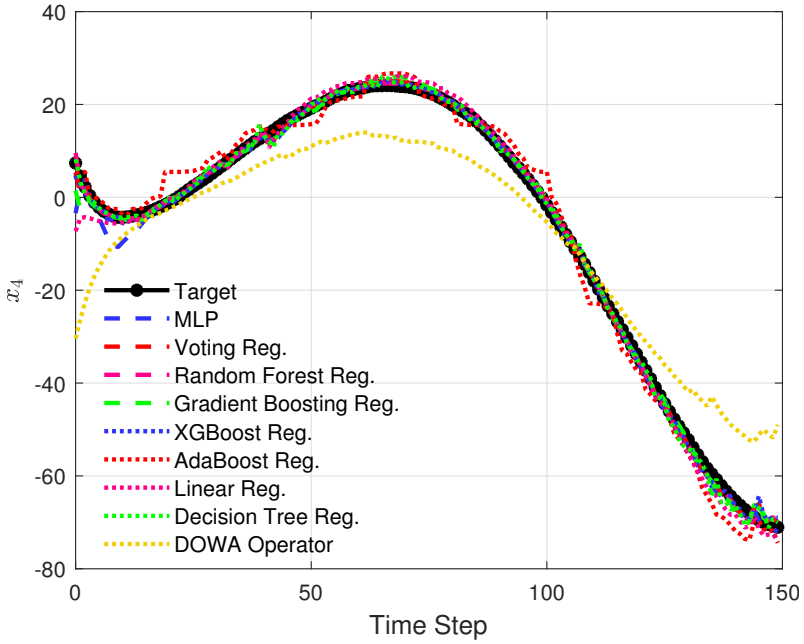

Figure 16. Fusion results for one Monte Carlo simulation for state 4.

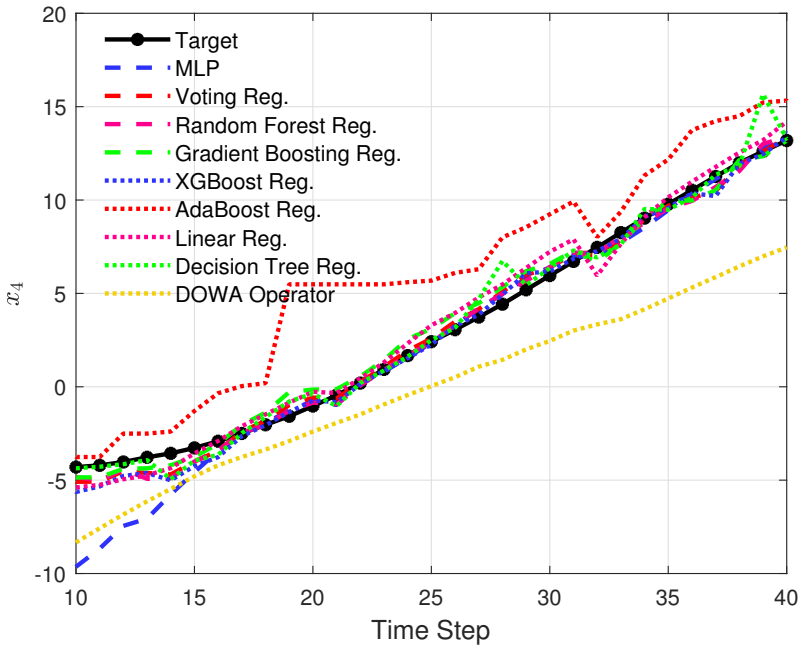

Figure 17. Fusion results for one Monte Carlo simulation for state 4 (zoomed in).

5. Conclusions

This paper explored the efficacy of state estimation fusion using machine learning (ML) regression methods on a linear smart microgrid based on an IEEE 4-bus model. For the very first time, to the best of the authors' knowledge, both cyberattack and packet loss for an unreliable IoT-based communication channel were considered simultaneously in this paper. The simulation results validate and justify the proposed idea, by considering the obtained RMSE, MAE and R-squared error indices, that the ML regression methods of interest are capable of functioning as data fusion methods, thus, improving the results of the KF estimators.

The aforementioned methods were employed to fuse the state estimates of four different data-sensing sources under the conditions of severe unreliability observed in the communication channel between the power substations and the energy management sys-

tem (EMS). The main advantage of using ML regression methods for data fusion is that they are relatively fast and accurate in terms of errors and predicting data trends. Their results are also easily explained and comprehended.

The results showed that ML regression methods can outperform MLP, which is a universal function approximator. This study also validated the inefficiency of the weighted averaging operators by leveraging the dependent ordered weighted averaging (DOWA) operator for the purpose of fusion.

Indeed, the hyperparameters of the ML regression methods are generally hard to tune, and some methods do not perform as expected in noisy environments. Nevertheless, these methods, combined with the proposed structure, are suitable candidates for dealing with cyberattacks and packet losses in smart microgrids. They can replace the complex and twisted methods, which demand high computational cost and complex mathematical formalization.

As for future works, the performance of other regression methods will be further investigated with a particular focus on multi-regression fusion methods, such as Decision Template, and the proposed method will be tested in other similar areas, such as cyber-physical systems, that involve sensor data estimation and sensor data fusion.

Author Contributions: Conceptualization, M.S., D.S.Z. and B.M.; methodology, M.S. and D.S.Z.; software, M.S., D.S.Z. and E.N.S.; validation, B.M., E.N.S., J.G.H. and J.M.M.L.; data curation, M.S. and D.S.Z.; writing—original draft preparation, M.S. and D.S.Z.; writing—review and editing, all authors; visualization, D.S.Z.; supervision, B.M.; funding acquisition, E.N.S., J.G.H. and J.M.M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by public research projects of Spanish Ministry of Science and Innovation, references PID2020-118249RB-C22 and PDC2021-121567-C22 - AEI/10.13039/501100011033, and by the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement with UC3M in the line of Excellence of University Professors, reference EPUC3M17.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Abbreviations

The following abbreviations are used in this manuscript:

DERs	distributed energy resources
ML	machine learning
EMS	energy management system
KF	Kalman filter
WSNs	wireless sensor networks
OWA	ordered weighted averaging
CKF	cubature Kalman filter
CPLs	constant power loads
SSRCKF	spherical simplex radial cubature Kalman filter
IoT	internet of things
WLS	weighted least squares
ML	maximum likelihood
KL	Kullback-Leibler
FDI	false data injection
TFMP	track-level fusion-based model prediction
KLPF	Kalman-like particle filter

R-ADMM	relaxed alternating direction method of multipliers
PGDL	physically-guided deep learning
DNNs	deep neural networks
ANN	artificial neural network
CDBN	conditional deep belief network
DSSE	distribution system state estimation
PCCs	points of common coupling
LTE	long-term evolution
BPSK	binary phase-shift keying
AWGN	additive white Gaussian noise
Log-MAP	log-maximum a posteriori
MLP	multi-layer perceptron
MRF	multi-regressor fusion
DOWA	dependent ordered weighted averaging
MC	Monte Carlo

References

1. Rodriguez-Diaz, E.; Vasquez, J.C.; Guerrero, J.M. Intelligent DC Homes in Future Sustainable Energy Systems: When efficiency and intelligence work together. *IEEE Consum. Electron. Mag.* **2016**, *5*, 74–80. [CrossRef]
2. Morstyn, T.; Hredzak, B.; Demetriades, G.D.; Agelidis, V.G. Unified Distributed Control for DC Microgrid Operating Modes. *IEEE Trans. Power Syst.* **2016**, *31*, 802–812. [CrossRef]
3. Haque, N.I.; Shahriar, M.H.; Dastgir, M.G.; Debnath, A.; Parvez, I.; Sarwat, A.; Rahman, M.A. Machine Learning in Generation, Detection, and Mitigation of Cyberattacks in Smart Grid: A Survey. *arXiv* **2020**, arXiv:2010.00661.
4. Sinopoli, B.; Schenato, L.; Franceschetti, M.; Poolla, K.; Jordan, M.I.; Sastry, S.S. Kalman filtering with intermittent observations. *IEEE Trans. Autom. Control.* **2004**, *49*, 1453–1464. [CrossRef]
5. Su, L.; Chesi, G. On the robust stability of uncertain discrete-time networked control systems over fading channels. In Proceedings of the 2015 American Control Conference (ACC), Chicago, IL, USA, 1–3 July 2015; pp. 6010–6015. [CrossRef]
6. Liao, W.; Salinas, S.; Li, M.; Li, P.; Loparo, K.A. Cascading Failure Attacks in the Power System: A Stochastic Game Perspective. *IEEE Internet Things J.* **2017**, *4*, 2247–2259. [CrossRef]
7. Esmalifalak, M.; Shi, G.; Han, Z.; Song, L. Bad Data Injection Attack and Defense in Electricity Market Using Game Theory Study. *IEEE Trans. Smart Grid* **2013**, *4*, 160–169. [CrossRef]
8. Rana, M.M.; Li, L. An Overview of Distributed Microgrid State Estimation and Control for Smart Grids. *Sensors* **2015**, *15*, 4302–4325. [CrossRef]
9. Singh, A.K.; Singh, R.; Pal, B.C. Stability Analysis of Networked Control in Smart Grids. *IEEE Trans. Smart Grid* **2015**, *6*, 381–390. [CrossRef]
10. Quevedo, D.E.; Ahlen, A. A predictive power control scheme for energy efficient state estimation via wireless sensor networks. In Proceedings of the 2008 47th IEEE Conference on Decision and Control, Cancun, Mexico, 9–11 December 2008; pp. 1103–1108. [CrossRef]
11. Chen, B.; Zhang, W.A.; Yu, L. Distributed Fusion Estimation With Missing Measurements, Random Transmission Delays and Packet Dropouts. *IEEE Trans. Autom. Control.* **2014**, *59*, 1961–1967. [CrossRef]
12. Kordestani, M.; Dehghani, M.; Moshiri, B.; Saif, M. A New Fusion Estimation Method for Multi-Rate Multi-Sensor Systems With Missing Measurements. *IEEE Access* **2020**, *8*, 47522–47532. [CrossRef]
13. Yin, Z.G.; Zhao, C.; Zhong, Y.R.; Liu, J. Research on Robust Performance of Speed-Sensorless Vector Control for the Induction Motor Using an Interfacing Multiple-Model Extended Kalman Filter. *IEEE Trans. Power Electron.* **2014**, *29*, 3011–3019. [CrossRef]
14. Niehsen, W. Information fusion based on fast covariance intersection filtering. In Proceedings of the Fifth International Conference on Information Fusion, FUSION 2002, (IEEE Cat. No. 02EX5997), Annapolis, MD, USA, 8–11 July 2002; Volume 2, pp. 901–904. [CrossRef]
15. Guo, Q.; Chen, S.; Leung, H.; Liu, S. Covariance intersection based image fusion technique with application to pansharpening in remote sensing. *Inf. Sci.* **2010**, *180*, 3434–3443. [CrossRef]
16. Rana, M.; Li, L.; Su, S. Distributed microgrid state estimation using smart grid communications. In Proceedings of the 2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Brisbane, QLD, Australia, 15–18 November 2015; pp. 1–5. [CrossRef]
17. Rana, M.M.; Bo, R.; Choi, B.J. Residual Saturation Based Kalman Filter for Smart Grid State Estimation Under Cyber Attacks. In Proceedings of the 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Suzhou, China, 29 July–2 August 2019; pp. 1459–1463. [CrossRef]
18. Kardan, M.A.; Asemani, M.H.; Khayatian, A.; Vafamand, N.; Khooban, M.H.; Dragicevic, T.; Blaabjerg, F. Improved Stabilization of Nonlinear DC Microgrids: Cubature Kalman Filter Approach. *IEEE Trans. Ind. Appl.* **2018**, *54*, 5104–5112. [CrossRef]

19. Kardan, M.A.; Moshiri, B.; Vafamand, N.; Razavi-Far, R.; Saif, M. Cyber Attack Estimation of Nonlinear DC Microgrids with Noisy Measurements: Spherical Simplex Radial CKF Approach. In Proceedings of the 2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I CPS Europe), Bari, Italy, 7–10 September 2021; pp. 1–6. [CrossRef]
20. De Sanctis, M.; Cianca, E.; Araniti, G.; Bisio, I.; Prasad, R. Satellite Communications Supporting Internet of Remote Things. *IEEE Internet Things J.* **2016**, *3*, 113–123. [CrossRef]
21. 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]
22. Alimardani, A.; Therrien, F.; Atanackovic, D.; Jatskevich, J.; Vaahedi, E. Distribution System State Estimation Based on Nonsynchronized Smart Meters. *IEEE Trans. Smart Grid* **2015**, *6*, 2919–2928. [CrossRef]
23. Meliopoulos, S.; Huang, R.; Polymeneas, E.; Cokkinides, G. Distributed dynamic state estimation: Fundamental building block for the smart grid. In Proceedings of the IEEE Power and Energy Society General Meeting, Denver, CO, USA, 26–30 July 2015; pp.1–6. [CrossRef]
24. Li, S.; Yilmaz, Y.; Wang, X. Quickest Detection of False Data Injection Attack in Wide-Area Smart Grids. *IEEE Trans. Smart Grid* **2015**, *6*, 2725–2735. [CrossRef]
25. Moslemi, R.; Mesbahi, A.; Velni, J.M. A Fast, Decentralized Covariance Selection-Based Approach to Detect Cyber Attacks in Smart Grids. *IEEE Trans. Smart Grid* **2018**, *9*, 4930–4941. [CrossRef]
26. Mustafa, A.; Poudel, B.; Bidram, A.; Modares, H. Detection and Mitigation of Data Manipulation Attacks in AC Microgrids. *IEEE Trans. Smart Grid* **2020**, *11*, 2588–2603. [CrossRef]
27. Rana, M.M.; Li, L.; Su, S.W. Cyber attack protection and control of microgrids. *Ieee/Caa J. Autom. Sin.* **2018**, *5*, 602–609. [CrossRef]
28. Khalid, H.M.; Peng, J.C. Immunity Toward Data-Injection Attacks Using Multisensor Track Fusion-Based Model Prediction. *IEEE Trans. Smart Grid* **2017**, *8*, 697–707. [CrossRef]
29. Rana, M.M.; Li, L.; Su, S.W. Distributed State Estimation of Smart Grids with Packet Losses. *Asian J. Control.* **2017**, *19*, 1306–1315. [CrossRef]
30. Rana, M.M.; Li, L.; Su, S.W.; Xiang, W. Consensus-Based Smart Grid State Estimation Algorithm. *IEEE Trans. Ind. Informat.* **2018**, *14*, 3368–3375. [CrossRef]
31. Zheng, W.; Li, Z.; Liang, X.; Zheng, J.; Wu, Q.H.; Hu, F. Decentralized State Estimation of Combined Heat and Power System Considering Communication Packet Loss. *J. Mod. Power Syst. Clean Energy* **2020**, *8*, 646–656. [CrossRef]
32. Qu, B.; Shen, B.; Shen, Y.; Li, Q. Dynamic state estimation for islanded microgrids with multiple fading measurements. *Neurocomputing* **2020**, *406*, 196–203. [CrossRef]
33. Noor-A-Rahim, M.; Khyam, M.O.; Mahmud, M.A.; Huque, M.T.I.U.; Li, X.; Pesch, D.; Oo, A.M. Robust and Real-Time State Estimation of Unstable Microgrids Over IoT Networks. *IEEE Syst. J.* **2021**, *15*, 2176–2185. [CrossRef]
34. Rana, M.M.; Bo, R.; Abdelhadi, A. Distributed Grid State Estimation under Cyber Attacks Using Optimal Filter and Bayesian Approach. *IEEE Syst. J.* **2021**, *15*, 1970–1978. [CrossRef]
35. Wang, L.; Zhou, Q.; Jin, S. Physics-guided Deep Learning for Power System State Estimation. *J. Mod. Power Syst. Clean Energy* **2020**, *8*, 607–615. [CrossRef]
36. Tanvir, A.A.; Merabet, A. Artificial Neural Network and Kalman Filter for Estimation and Control in Standalone Induction Generator Wind Energy DC Microgrid. *Energies* **2020**, *13*, 1743. [CrossRef]
37. Bhusal, N.; Shukla, R.M.; Gautam, M.; Benidris, M.; Sengupta, S. Deep ensemble learning-based approach to real-time power system state estimation. *Int. J. Electr. Power Energy Syst.* **2021**, *129*, 106806. [CrossRef]
38. He, Y.; Mendis, G.J.; Wei, J. Real-Time Detection of False Data Injection Attacks in Smart Grid: A Deep Learning-Based Intelligent Mechanism. *IEEE Trans. Smart Grid* **2017**, *8*, 2505–2516. [CrossRef]
39. Deng, R.; Zhuang, P.; Liang, H. False Data Injection Attacks Against State Estimation in Power Distribution Systems. *IEEE Trans. Smart Grid* **2019**, *10*, 2871–2881. [CrossRef]
40. Li, H.; Li, F.; Xu, Y.; Rizy, D.T.; Kueck, J.D. Adaptive Voltage Control With Distributed Energy Resources: Algorithm, Theoretical Analysis, Simulation, and Field Test Verification. *IEEE Trans. Power Syst.* **2010**, *25*, 1638–1647. [CrossRef]
41. Li, H.; Lai, L.; Poor, H.V. Multicast Routing for Decentralized Control of Cyber Physical Systems with an Application in Smart Grid. *IEEE J. Sel. Areas Commun.* **2012**, *30*, 1097–1107. [CrossRef]
42. Rana, M.M.; Li, L. Kalman Filter Based Microgrid State Estimation Using the Internet of Things Communication Network. In Proceedings of the 2015 12th International Conference on Information Technology—New Generations, Las Vegas, NV, USA, 13–15 April 2015; pp. 501–505. [CrossRef]
43. Liu, Y.; Ning, P.; Reiter, M.K. False data injection attacks against state estimation in electric power grids. In Proceedings of the 16th ACM Conference on Computer and Communications Security,Chicago, IL, USA, 9–13 November 2009; pp. 21–32. [CrossRef]
44. Yuan, Y.; Li, Z.; Ren, K. Modeling Load Redistribution Attacks in Power Systems. *IEEE Trans. Smart Grid* **2011**, *2*, 382–390. [CrossRef]
45. Zhang, J.; Chu, Z.; Sankar, L.; Kosut, O. False data injection attacks on phasor measurements that bypass low-rank decomposition. In Proceedings of the 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm), Dresden, Germany, 23–27 October 2017; pp. 96–101. [CrossRef]

46. Deshmukh, S.; Natarajan, B.; Pahwa, A. State Estimation Over a Lossy Network in Spatially Distributed Cyber-Physical Systems. *IEEE Trans. Signal Process.* **2014**, *62*, 3911–3923. [CrossRef]
47. Jiang, Y. *A Practical Guide to Error-Control Coding Using MATLAB*, 1st ed.; Artech House Publishers: London, UK 2010; p. 281.
48. Rana, M.M.; Li, L.; Su, S.W. Distributed condition monitoring of renewable microgrids using adaptive-then-combine algorithm. In Proceedings of the 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, USA, 17–21 July 2016; pp. 1–5. [CrossRef]
49. Simon, D. *Optimal State Estimation*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006; pp. 1–526. [CrossRef]
50. Baltrusaitis, T.; Ahuja, C.; Morency, L.P. Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 423–443. [CrossRef]
51. Meng, T.; Jing, X.; Yan, Z.; Pedrycz, W. A survey on machine learning for data fusion. *Inf. Fusion* **2020**, *57*, 115–129. [CrossRef]
52. Brena, R.F.; Aguileta, A.A.; Trejo, L.A.; Molino-Minero-Re, E.; Mayora, O. Choosing the Best Sensor Fusion Method: A Machine-Learning Approach. *Sensors* **2020**, *20*, 2350. [CrossRef]
53. Moshiri, B. Tutorial A: Sensor data fusion, principles and applications. In Proceedings of the 2010 International Symposium on Optomechatronic Technologies, Toronto, ON, Canada, 25–27 October 2010; pp. 1–2. [CrossRef]
54. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
55. Mason, L.; Baxter, J.; Bartlett, P.; Frean, M. Boosting Algorithms as Gradient Descent. In Proceedings of the 12th International Conference on Neural Information Processing Systems, Denvor, CO, USA, 29 November–4 December 1999; pp. 512–518.
56. Hastie, T.; Tibshirani, R.; Friedman, J. Boosting and Additive Trees. In *The Elements of Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 337–387.
57. Xu, Z. Dependent OWA Operators. In *Modeling Decisions for Artificial Intelligence*; Torra, V., Narukawa, Y., Valls, A., Domingo-Ferrer, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; Volume 3885, pp. 172–178. [CrossRef]
58. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [CrossRef]
59. Soleymannejad, M.; Basiri, A. Using OWA Approach to Solve Cold-Start Problem of Recommender Systems. In Proceedings of the 2020 10th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 29–30 October 2020; pp. 500–507. [CrossRef]
60. Kazemi, E.; Sadrian Zadeh, D.; Moshiri, B. Metal-oxide-semiconductor Sensors Modeling Using Ordered Weighted Averaging (OWA) Operators in Electronic Nose. *Measurement* **2021**, *184*, 109932. [CrossRef]