IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 2, 2021, accepted June 17, 2021, date of publication June 25, 2021, date of current version July 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3092308

State Estimation for Hybrid VSC Based HVDC/AC: Unified Bad Data Detection Integrated With Gaussian Mixture Model

MOTAZ AYIAD[1,2], HELDER LEITE[2], AND HUGO MARTINS[1]
[1]Efacec Automation, Grid Management Division, 4471-907 Porto, Portugal
[2]Faculty of Engineering (FEUP), University of Porto, 4200-465 Porto, Portugal

Corresponding author: Motaz Ayiad (motaz.ayiad@efacec.com)

This work was supported by the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant Agreement 765585. This document reflects only the author's views; the European Commission is not responsible for any use that may be made of the information it contains.

ABSTRACT The High Voltage Direct Current (HVDC) is an emerging technology that transmits power over long distances and at a higher capacity than the traditional AC systems. Integration of HVDC into modern power networks requires vital modification to the Supervisory, Control and Data Acquisition (SCADA) system, particularly in power system applications. For instance, the state estimator toolbox is an essential software to estimate the network AC and DC systems states. However, the state estimator may fail due to severely corrupted data, a.k.a bad data; hence, an additional data treatment is needed. This paper presents a unified bad data detection block for Weighted Least Squares (WLS) state estimation algorithm. The bad data detection block works for hybrid Voltage Source Converter (VSC)-HVDC/AC transmission systems. It improves the traditional Largest Normalized Residual (LNR) method by integrating the Gaussian Mixture Model (GMM) algorithm. This method reduces the time needed for bad data detection, increases the algorithm robustness, and enhances estimation accuracy. The Cigre B4 network is used as a test case to validate this work on a hybrid VSC-HVDC/AC network. Also, grid load and generation data from the UK is used to construct the simulation measurements and the GMM model. Simulation results show that the modified GMM-LNR has considerably reduced the detection time and improved the WLS accuracy.

INDEX TERMS State estimation, VSC, HVDC, AC/DC, bad data, unified, GMM, LNR.

I. INTRODUCTION

Modern power networks are becoming more complex due to the penetration of large-scale low-carbon generation units such as solar and wind farms. In long-distance high power transmission, the power is transferred through a modern grid technology known as HVDC grids [1]. The HVDC has reduced the overall power transmission losses and provided better performance over the conventional AC networks [1]–[3]. Besides, it has allowed interconnecting asynchronous networks and transfers high power capability [2], [4], [5]. The VSC based HVDC provide high controllability over the generated power by independently control the active and reactive power. The architecture of the modern VSC-HVDC/AC stations and transmission grids are commonly operated by separate SCADA systems, as shown in Fig. 1 [1], [3], [5]–[7].

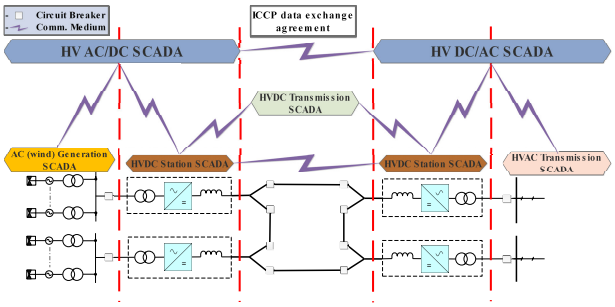


FIGURE 1. Modern SCADA systems architecture for hybrid HVDC/AC network [8].

The hybrid VSC-HVDC/AC networks are increasing and alongside that, the demand on more robust control of such grids becomes more essential and critical [1], [3], [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Min Wang.

The operators of the AC and DC networks are monitoring the grids through either centralized or distributed SCADA systems. In the SCADA command center, several power system toolboxes exist to control and observe the power network. The power system state estimation is one of the major tools in the layer of power system applications. It processes the imperfect (noisy) and redundant measurements to provide more accurate presentation of the power network states (i.e. DC voltages and AC voltages and phase angles) [9], [10]. These measurements are collected by data acquisition devices represented by Remote Terminal Units (Phasor Measurements Unit (PMU) and Intelligent Electronic Devices (IED)) [11].

Research has been chunked with methods and modifications for transmission network state estimation algorithms, especially on the AC side. The most known algorithm, the WLS, was first formulated by Fred Schweppe in 1969 [12], it is still used on transmission level alongside it's common competitor, the Kalman filter [13]. Several other methods have been proposed in the past years; some focused on the computational demand, such as fast decoupled state estimation, while others on the accuracy and bad data detection (i.e. robustness) such as Hachtel's Augmented Matrix (HAM), exact second-order state estimation and Least Absolute Value (LAV). A comparison conducted in [14] has concluded that WLS is performing better than other methods in cases of noisy measurements and zero bad data.

The integration of the VSC-HVDC grids has raised challenges to the state estimator toolbox and affected it's performance, efficiency and reliability [15]–[17]. As a result, there is an ongoing research on the state estimation toolbox for HVDC, and mainly on hybrid VSC-HVDC/AC networks. The fast decoupled state estimator for hybrid AC/DC systems is presented in [18]. In [17], a robust state estimation for VSC-DC systems based on LAV method is proposed. While in [19] and [20], sequential weighted state estimator is provided for multi-terminal VSC-DC/AC grids. The PMUs based state estimators are discussed in [21] and [22]. HAM and WLS algorithms are used in [23] and [8] respectively as a unified state estimator for VSC-HVDC/AC, both argued that a unified (centralized) approach increase the estimation accuracy due to the higher redundancy rate of measurements.

Despite that high-end devices (PMUs) are widely installed in transmission networks, the possibility of having bad data still exists. Faulty devices, environmental conditions, timestamps mismatch and cyber-attacks are the common reasons for these defected data [24]. Modern data-driven algorithms have emerged recently, a deep learning bad data detection algorithm for linear state estimator is available in [25], and for AC smart-grids in [26]. A machine learning approach based on PMU is proposed in [27]. Conventional algorithms with internal bad data rejection such as LAV is presented in [17], [28], and with HAM in [23], but these methods can be more computational demanding. In the traditional WLS, the bad data detection requires an additional post-processing block, which increases the computational processing. Roughly, each single bad data requires a complete WLS state estimation cycle. The most common bad data methods for WLS is the Chi-square and LNR approach [17], [29]. The work in [30] proposes a mixture Gaussian distribution learning method to detect the false data injection attacks on smart grids state estimations.

Regardless of the variety of state estimation and bad data detection algorithms, the WLS method is still commonly used in the SCADA of transmission networks. Therefore, the work in this paper focuses on keeping the WLS algorithm intact while improving its robustness by integrating the GMM with the LNR bad data detection block. The approach requires historical data profiles to be modelled using GMM, since it has shown good performance in load profile modelling [31]–[33]. Then, the LNR algorithm is modified to use the GMM parameters since the LNR is a function of the measurement's weights. GMM provides estimates of the variances of the measurement by clustering the data profiles into groups. The estimated variances are used to better weigh the measurements during the state estimation and the bad data detection processes. The measurements that appear outside the clustered groups are in advance suspected to be bad data.

The proposed approach has achieved several contributions: (1) The method is compatible and deployable for VSC-HVDC/AC unified state estimation. Furthermore, the additional blocks require no changes on the WLS algorithm or the traditional LNR, which builds interests for the industrial vendors. (2) The pre-processing block calculates the GMM parameters, passes them to the WLS block, and identifies the suspected bad data in advance. (3) The post-processing GMM-LNR block checks the suspected data and corrects them based on the GMM clusters. (4) The computational efforts have been reduced compared with the traditional LNR bad data detector. In short, these modifications have increased the state estimation accuracy and reduced the processing time of bad data detection.

The structure of the paper is as follows: Section II reviews the WLS state estimator and the traditional bad data detection. Section III goes through the mathematical background of the GMM and the integration of it with WLS. Section IV presents the results and the validation of the modified method. Section V concludes the work of this paper.

II. STATE ESTIMATION AND BAD DATA DETECTION

The state estimator toolbox in the SCADA system provides the most accurate network states (voltages and angles) possible. It processes the collected measurements (z) from the network by reducing noise (error e) added to the measurements during the acquisition and transferring of data as shown in the equation below:

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

The measurements usually come from three sources:

1) PMUs or high-end devices: these devices provide very accurate (low uncertainty) measurements to the

state estimator. Usually, these measurements have a variance between 10^{-6} to 10^{-8}.

2) IEDs: measurement and control devices with an accuracy that variates depending on the purpose and location. It's variance can vary between 10^{-3} and 10^{-6}. The measurements from these devices are commonly assumed to have Gaussian noise.
3) Pseudo-Measurements: the source of these measurements is the historical data and statistical models or more advanced neural networks (machine learning today). The statistical model measurements have high uncertainty with a variance between 10^{-1} and 10^{-2}. These measurement are usually used in cases of an unobserved network (not enough measurements).

A. WEIGHTED LEAST SQUARES (WLS)

The WLS algorithm is one of the common tools available today to provide better system states observation. The algorithm's objective function $J(x)$ is represented by the square error (residual r) of the measurements and their weights. The WLS can be formulated as a matrices minimization problem as shown in (2).

$$\underset{x}{\text{Minimise}}: \; J(x) = (z - h(x))^T R^{-1} (z - h(x)) \tag{2}$$

where, $z = \left[z_1 \; z_2 \; \ldots \; z_n\right]^T$ is the measurements vector with n measurements, x is the state variable vector, $h(x) = \left[h_1(x) \; h_2(x) \; \ldots \; h_n(x)\right]^T$ is a vector of non-linear functions evaluated at the state variable x, and R is a diagonal matrix that contains the variances (σ^2) of all measurements, $R = \text{diag}\left[\sigma_1^2 \; \sigma_2^2 \; \ldots \; \sigma_n^2\right]^T$ for independent measurements. The WLS state estimation steps are illustrated in Fig. 2 flowchart.

Using the Taylor series approximation and Newton-Raphson, the minimization problem can be simplified to (3)

$$\hat{x}_{k+1} = \hat{x}_k + \Delta\hat{x}_k \tag{3}$$

where $\Delta\hat{x}_k$ is the system states update factor, and calculated as follow:

$$\Delta\hat{x}_k = \left[G(\hat{x}_k)\right]^{-1} \nabla F(\hat{x}_k) \tag{4}$$

$$\nabla F(\hat{x}_k) = H^T(x_k) R^{-1}(x_k)(z - h(x_k)) \tag{5}$$

$$G(x_k) = H^T(x_k) R^{-1}(x_k) H(x_k) \tag{6}$$

where $G(x_k)$ is the Gain matrix and $H(x) = \frac{\partial h(x)}{\partial x}$ is the Jacobian matrix.

In a hybrid VSC-HVDC/AC system, the state variable vector x contains AC voltages and phase angles and DC voltages. Therefore, the unified WLS measurement functions vector $h(x)$ is extended to include AC, DC and converter measurements and constraints as shown below:

$$h(x) = \begin{bmatrix} h_{AC_1}(x) & h_{AC_2}(x) & \ldots & h_{AC_a}(x) \\ h_{DC_1}(x) & h_{DC_2}(x) & \ldots & h_{DC_b}(x) \\ hpc_1(x) & hpc_2(x) & \ldots & hpc_c(x) \\ hvc_1(x) & hvc_2(x) & \ldots & hvc_c(x) \end{bmatrix}^T$$

where $h_{AC}(x)$ and $h_{DC}(x)$ are the AC and DC measurements sets respectively, the $hpc(x)$ and $hvc(x)$ represent the converter power constraints and the voltage coupling measurements, and a, b and c are the number of AC systems, DC systems and converters respectively.

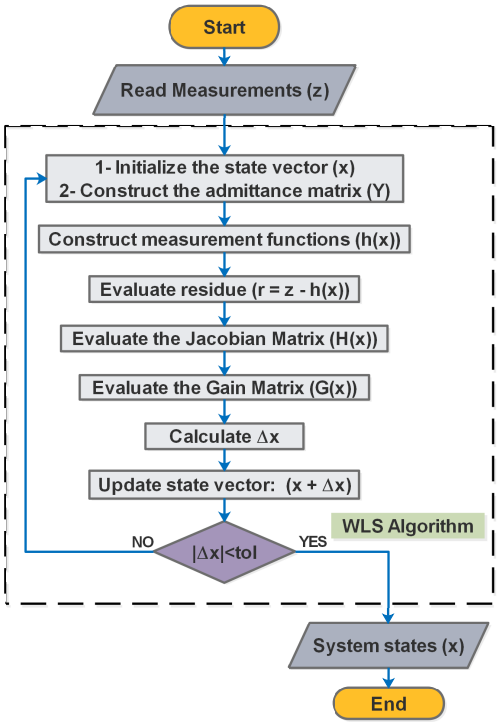


FIGURE 2. The WLS state estimation algorithm.

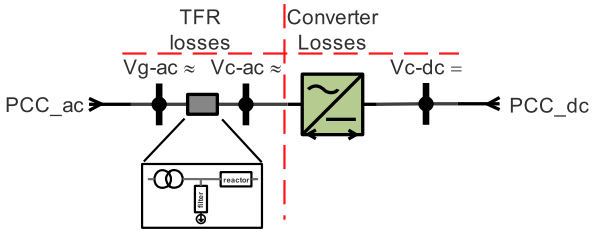


FIGURE 3. Converter power coupling schematic [8].

The converter Power Coupling ($hpc(x)$) is presented as a zero-constraint function between the AC/DC power and the converter losses as shown in (7) [8], [34].

$$hpc_j(x_{ac}, x_{dc}) = P_{ac}(x_{ac}) + P_{dc}(x_{dc}) + P_{cl} = 0 \tag{7}$$

where x_{ac} is the voltage and phase angle states of the AC side, x_{dc} is the DC side voltage state, j is the converter number, and P_{cl} is the converter losses. P_{cl} can be calculated using an experimental statistical approach available in [4], [8], [10]. In Fig. 3, the TFR term presents the AC side transformer, filter and reactor losses.

The converter Voltage Coupling ($hvc(x)$) is presented as a ratio measurement M_{idx} (modulation index) between the AC and DC sides voltages as shown in Fig. 4. M_{idx} is calculated using (8) [8], [17], [35].

$$hvc_j(x_{ac}, x_{dc}) = M_{idx} = K_{cf} \frac{V_{c-dc}}{V_{g-ac}} \tag{8}$$

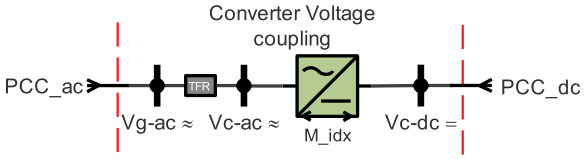


FIGURE 4. Converter voltage coupling schematic [8].

where x_{ac} and x_{dc} are the voltage states of the AC and DC sides for the jth converter, V_{g-ac} is related to the converter operation conditions (rectifier/inverter), and K_{cf} is the voltage conversion factor, it can have different values based on the AC side network topology and converter type. Further details are available [8].

In addition, the Jacobian matrix is modified to cover the extensions as shown the matrix below $H_{unified}$.

$$H_{unified}\,(x) = \begin{bmatrix} H_{AC-AC} & H_{DC-AC} \\ H_{AC-DC} & H_{DC-DC} \\ H_{hpc-AC} & H_{hpc-DC} \\ H_{hvc-AC} & H_{hvc-DC} \end{bmatrix}$$

where

H_{AC-AC} is the $\partial h_{AC}(x)$ to θ and V_{ac}
H_{DC-AC} is the $\partial h_{DC}(x)$ to θ and $V_{ac} \leftrightarrow$ 0-matrix
H_{AC-DC} is the $\partial h_{AC}(x)$ to $V_{dc} \leftrightarrow$ 0-matrix
H_{DC-DC} is the $\partial h_{DC}(x)$ to V_{dc}
H_{hpc-AC} is the $\partial hpc(x)$ to θ and V_{ac}
H_{hpc-DC} is the $\partial hpc(x)$ to V_{dc}
H_{hvc-AC} is the $\partial hvc(x)$ to θ and V_{ac}
H_{hvc-DC} is the $\partial hvc(x)$ to V_{dc}

The Jacobian matrix can be expanded as shown below, for a AC systems, b DC Systems, and c converters $H(x)$, as shown at the bottom of the page.

Further details on the mathematical background and the equations of the unified VSC-HVDC/AC state estimator are available in [8].

B. UNIFIED BAD DATA DETECTION

Bad Data processing aims to detect, identify and eliminate/correct measurements with gross errors. These measurements are usually corrupted with non-Gaussian errors. In an over-observed network, the bad data detection process becomes easier due to the effect of the redundant measurements. It is essential to avoid loss of observability whenever a measurement elimination is required [23].

Gross errors usually occur due to device failure, cyber-attacks or out-of-synchronization communication. In WLS, there are two common techniques used to detect bad data, the χ^2 test and the LNR.

1) CHI-SQUARE (χ^2) TEST

The χ^2 test is a statistical method based on chi-square distribution to detect bad data [36]. The test compares the state estimation objective function with the χ^2 distribution. The χ^2 value is evaluated at $(m - s)$ degree of freedom and $C\%$ confidence interval, where m is the number measurement, s is the number of states, and $C\%$ is chosen by the SCADA operator and commonly 95% is used.

This method detects the existence of bad data if $J(x)$ evaluated by (9) is greater than the estimated χ^2, where r_i is the residual, calculated by $z_i - h_i(x)$.

$$J(x) = \sum_{i=1}^{m} \frac{{r_i}^2}{R_{ii}} \tag{9}$$

The χ^2 test is not computationally demanding, but it is not sufficient to identify the bad measurement on its own.

2) LARGEST NORMALIZED RESIDUAL (LNR) TEST

The LNR is a hypothesis testing method used to identify the bad measurement. It starts after the doubts of bad data existence from the chi-square test. The integration of LNR in WLS state estimation algorithm is shown in Fig. 5, and its processes can be described as follow:

1) Check if bad data exists in the measurements set (e.g. using χ^2 test);
2) Compute the LNR for all measurements based on the below equation:

$$r_i^{NLR} = \frac{|r_i|}{\sqrt{S_{ii}}} \tag{10}$$

where S is the residual sensitive matrix given by:

$$S = R - (H(H^T R^{-1} H)^{-1} H^T) \tag{11}$$

$$H(x) = \left[\begin{array}{ccc|ccc} H_{(AC-AC)_1} & 0 & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & H_{(AC-AC)_a} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & H_{(DC-DC)_1} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & H_{(DC-DC)_b} \\ \hline H_{(hpc-AC)_1} & 0 & 0 & H_{(hpc-DC)_1} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & H_{(hpc-AC)_c} & 0 & 0 & H_{(hpc-DC)_c} \\ \hline H_{(hvc-AC)_1} & 0 & 0 & H_{(hvc-DC)_1} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & H_{(hvc-AC)_c} & 0 & 0 & H_{(hvc-DC)_c} \end{array}\right]$$

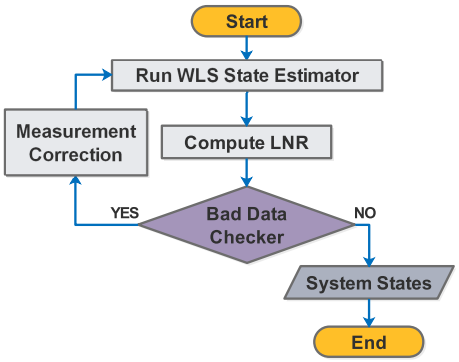


FIGURE 5. WLS with LNR bad data block.

3) Define a threshold *thr* which identifies the suspected measurements by applying the following if-statement:

$$\text{For all } r^{NLR} \begin{cases} \text{if } r_i^{NLR} > thr, & \text{Suspected} \\ \text{if } r_i^{NLR} \leq thr, & \text{Pass} \end{cases} \quad (12)$$

The *thr* value varies between 1.5 and 4 based on several papers and industrial reports, but the value 3 is commonly used [28].

4) The $max(r_i^{NLR})$ of all suspected measurements is selected to be eliminated or corrected.
5) Restart the process from the state estimator block with the updated measurements set.

The LNR method faces several drawbacks:

1) The method is considered accurate in detecting a single bad data at a time. However, it requires a complete restart of the WLS cycle per bad data. Furthermore, it finds difficulties with interacting and conforming bad measurements.
2) It increases the WLS computational burden due to the complex evaluation of the residual sensitive matrix.
3) The algorithm cannot avoid loss of observability when the measurement is removed. Therefore, additional conditions are required to maintain observability.

The process of eliminating a bad data simply means removing it from the measurement set. However, another approach can be implemented to correct the bad measurement using the equation below [28]:

$$z_{NLR-corrected} = z_{BD} - \frac{R_{ii}}{S_{ii}} \times r_i \quad (13)$$

The bad data block is essential to guarantee the robustness of the WLS algorithm whenever the integrity of measurement devices is not ensured. In this work, the LNR is expanded to deal with AC, and DC residuals and reinforced by GMM.

III. GAUSSIAN MIXTURE MODEL (GMM)

The GMM is a superposition statistical structure based on Gaussian distribution (mixtures/clusters). It clusters the data intro several K Gaussian distributions, each with a different parameters: mean (μ), variance (σ^2) and weight (fraction $\emptyset$). The GMM has a probability density function (14), and it can

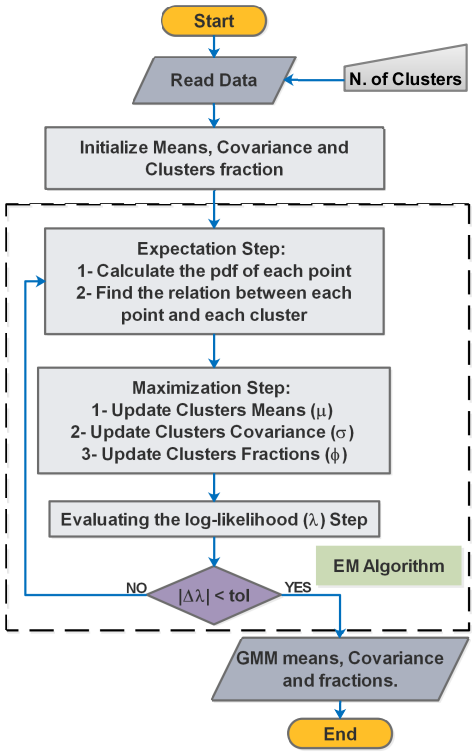


FIGURE 6. GMM and EM algorithm flowchart.

be derived by following the flowchart steps in Fig. 6. The Expectation-Maximization (EM) algorithm is used to obtain the GMM parameters [37].

$$GMM_{pdf} = \sum_{j=1}^{K} \emptyset_i G_{pdf} = \sum_{j=1}^{K} \emptyset_j G\left(\mu_j, \sigma_j^2\right) \quad (14)$$

GMM can be applied to 1-dimension of data as in (15) or d-dimensions as in (16) [31], where z is the data sample, σ is the variance, $\mathbb{E}$ is the covariance matrix in case of multivariate Gaussian distribution, and $\Gamma = \left(-\frac{1}{2}(z-\mu)^T \mathbb{E}^{-1}(z-\mu)\right)$.

$$G\left(z|\mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(z-\mu)^2\right) \quad (15)$$

$$G\left(z|\mu, \mathbb{E}\right) = \frac{1}{\left(\sqrt{2\pi}\right)^d \sqrt{\det(\mathbb{E})}} \exp(\Gamma) \quad (16)$$

A. THE EXPECTATION-MAXIMIZATION ALGORITHM

EM is one of the elegant iterative techniques to obtain the GMM parameters; it is used to cluster data [37]. The conventional EM process can be summarized as follow:

1) The Initialization step initializes the:
   - Means (μ): are given random values from the data set or can be obtained using K-means algorithm
   - Covariance ($\sigma/\mathbb{E}$): is a $D \times K$ matrix, where D is the data dimensions and K is the number of clusters. The covariance of all clusters in each dimension is initialized with the variance of that dimension.

- Clusters fractions ($\emptyset$): are initialized equally by assuming that each cluster shares $\frac{1}{K}$ of weight.

2) The Expectation step calculates the pdf of each data point (x) using the Gaussian pdf equation (15) or (16). Then finds the probability $f(j|z)$ that each data point belongs to a specific cluster j using (17), where K is the clusters number and $G_{pdf(j|z_i)} = z_i \times \mathcal{N}(\mu_j, \sigma_j^2)$.

$$f(j|z_i) = \frac{G_{pdf(j|z_i)} \times \emptyset_j}{\sum_{k=1}^{K} (G_{pdf(k|z_i)} \times \emptyset_k)} \tag{17}$$

3) The Maximization step takes the output of the Expectation step to update the clusters means (18), covariances (19), and fractions (weights) (20), using the weighted average equation, where N is the data length.

$$\mu_j = \frac{\sum_{i=1}^{N} f(j|z_i) \times z_i}{\sum_{i=1}^{N} f(j|z_i)} \tag{18}$$

$$\mathbb{E} = \frac{\sum_{i=1}^{N} f(j|z_i) \times (z_i - \mu_j) \times (z_i - \mu_j)^T}{\sum_{i=1}^{N} f(j|z_i)} \tag{19}$$

$$\emptyset_j = \frac{1}{N} \sum_{i=1}^{N} f(j|z_i) \tag{20}$$

4) The Termination step, the EM steps are iterated till either: 1- exceed the allowed number of iterations: 2- reaching an acceptable tolerance $\Delta\mu < tol$.

The outputs of GMM and EM are means, variances and clusters fractions; each mean and variance are corresponding to a specific cluster (Gaussian distribution pdf). The above procedure is summarised in the flowchart, Fig. 6.

The GMM probability density function GMM_{pdf} is the summation of the Gaussian mixtures, and for a single dimension data is presented (21).

$$GMM_{pdf} = \sum_{j=1}^{K} \emptyset_j \times G_{pdf(j|z_i)} \tag{21}$$

However, the traditional EM mentioned above has some convergence issues, and weak termination condition. Therefore, the problem is transformed into a maximization log-likelihood optimization problem (λ_{GMM}) as shown in (22), where N is the data length [37].

$$\lambda_{GMM} = \sum_{i=1}^{N} \log \left(\sum_{j=1}^{K} \emptyset_j \times G_{pdf(j|z_i)} \right) \tag{22}$$

The equations of means, covariance and weights are transformed to logarithmic form, e.g. (17) becomes:

$$\log f(j|z_i) = \log \left(G_{pdf(j|z_i)} \times \emptyset_j \right) - \log \sum_{k=1}^{K} G_{pdf(k|z_i)} \times \emptyset_k \tag{23}$$

In the expansion of the previous equation, the part $\log \sum_1^k G_{pdf} \times \emptyset$ is a log-sum-exp problem, and it requires a mathematical trick to be evaluated correctly without arithmetic underflow [38].

The modified version of the EM has a log-likelihood convergence, which offers a better termination condition ($\Delta\lambda_{GMM} < tol$). Further modification can be made on the conventional GMM to merge the mixtures and reduce the clusters number, this approach known as Reduced Gaussian Mixture Model (RGMM).

B. REDUCED GAUSSIAN MIXTURE MODEL (RGMM)

The RGMM aims to improve the accuracy of the GMM pdf approximation by combining some of the Gaussian mixtures together [31]. RGMM requires updating the weights, means and variances of the merged clusters using (24), (25) and (26), where j refers to the mixture to combine, and $\mathbb{E}$ is the covariance matrix.

The reduction of the clusters leads to loss in the total weights, i.e. $\sum_{i=1}^{K} \emptyset_i \neq 1$. Therefore, the mean and variances of the merged cluster is normalised with $\emptyset_{new}$.

$$\emptyset_{new} = \sum_{j \in I} \emptyset_j \tag{24}$$

$$\mu_{new} = \frac{1}{\emptyset_{new}} \sum_{j \in I} \emptyset \times \mu_j \tag{25}$$

$$\mathbb{E}_{new} = \frac{1}{\emptyset_{new}} \sum_{j \in I} \emptyset_j \times \left[\mathbb{E}_j + \xi \right]$$

$$\text{where } \xi = \left(\mu_j - \mu_{new} \right) \left(\mu_j - \mu_{new} \right)^T \tag{26}$$

RGMM algorithm starts by selecting the cluster with the largest weight as a principle cluster p, then merging any cluster j stands within a certain distance $D_{p->j}$. The distance is evaluated using the Salmond equation (27) [39], and compared with χ^2 threshold at a specific confidence level (e.g. 99%).

$$D_{p->j} = \frac{\emptyset_j \times \emptyset_p}{\emptyset_j + \emptyset_p} \left(\mu_p - \mu_j \right)^T \times \mathbb{E}_p^{-1} \times \left(\mu_p - \mu_j \right) \tag{27}$$

For all clusters, if $D_{p->j} < thr$ then cluster j is merged with the principle cluster p.

C. GMM MODIFICATION OVER THE TRADITIONAL LNR

The additional GMM data pre-processing block is shown in Fig. 7 flowchart. It shows the procedure of obtaining the GMM parameters required for the WLS algorithm. The GMM provides means and variances to the WLS to tune the estimation process more accurately. The bad data block determines the suspected bad data measurement by checking all measurements with the following constraint: for measurement z_i and for all clusters from 1 to j, z_i is a suspected measurement $\iff |z_i - \mu_j| > 3\sigma_j$.

After determining the suspected data, the GMM cluster j that represents the current load form is identified. Then, the calculations of the sensitivity matrix and the z_{NLR} are evaluated for the suspected elements using (13) and the

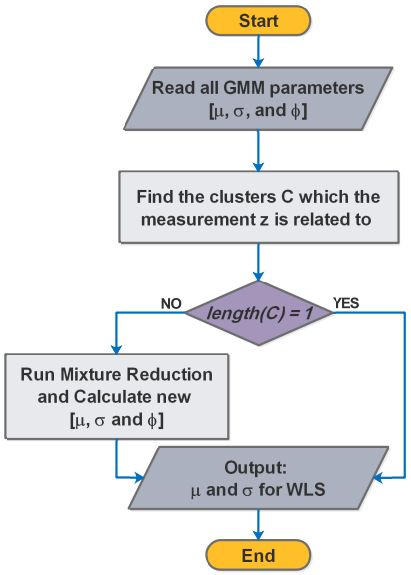


FIGURE 7. The GMM parameters for WLS flowchart.

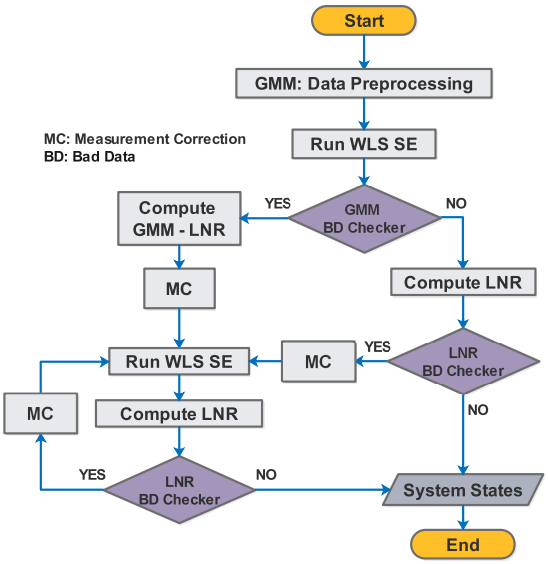


FIGURE 8. The GMM modification on WLS BD detection.

GMM variances. A GMM_{factor} is generated as a pseudo measurement, where $GMM_{factor} = \mathcal{N}(\mu_{j-GMM}, 3\sigma^2_{j-GMM})$.

If the z_{NLR} measurement is larger than the GMM_{factor}, then the measurement is substituted by a quasi-pseudo measurement using (28), if not then the z_{NLR} measurement is returned. The integration of GMM with LNR and WLS is shown in Fig. 8.

$$z_{GMM-NLR} = \frac{z_{NLR} \times GMM_{factor}}{z_{NLR} + GMM_{factor}} \tag{28}$$

IV. SIMULATION AND OUTCOMES

A. NETWORK DESCRIPTION

The network in Fig. 9 shows a modified version of the Cigre B4 HVDC/AC grid [40]. The network is assumed to cover part of the UK total demand at four demand points as shown

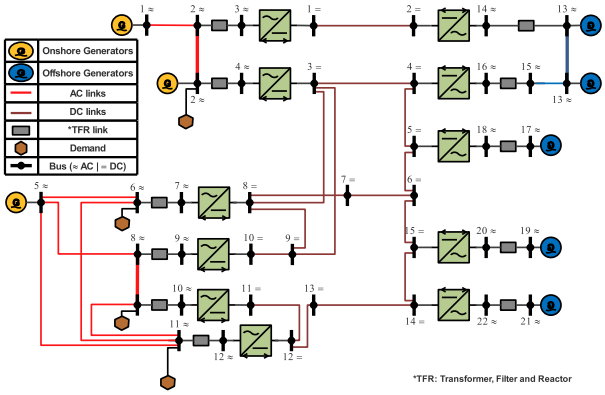


FIGURE 9. The Cigre B4 HVDC/AC transmission network.

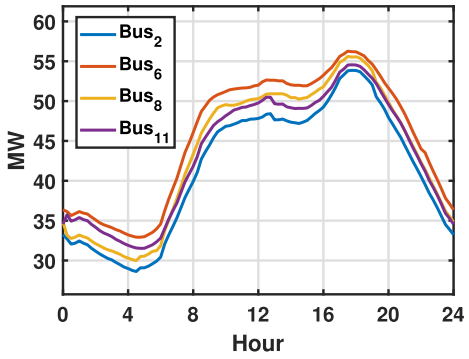


FIGURE 10. AC demand (load profile) of bus 2, 6, 8 and 11.

in the figure. The load and generation profiles are obtained from the UK NationalGrid historical data for January 2021 [41]. A simple data treatment has been carried out as follow:

1) The resolution of load profiles data is upscaled from half an hour to 15 minutes (48 to 96 data points per day). 14 days of data are aggregated into 7 days, then the moving mean algorithm is applied. Fig. 10 shows the upscaled load profiles of the demand buses.
2) The generation data are selected from the same period. 8 different days of wind generation profiles are chosen, one per unit generation in the Cigre B4. The data are downscaled from half an hour to one hour. Hence, the generation is assumed fixed for one hour.

B. LOAD PROFILE MODELLING USING GMM

The subfigures in Fig. 11 show the histogram and the GMM pdfs for the load profiles of bus 2, 6, 8 and 11. The width of the histogram bars is calculated using Scott's Rule [42], and the GMM is constructed with 3 clusters.

1) THE CHI SQUARE GOODNESS-OF-FIT TEST

The goodness-of-fit test is based on Chi-Square statistical test, and is used to test the performance of the GMM against other known distributions [43]. Figs. 12 and 13 show the goodness-of-fit test for Normal, Lognormal Gamma, Weibull and GMM distributions at load bus 2, and 11. The lower the Chi-Square statistics value, the better fit. There is a trade-off

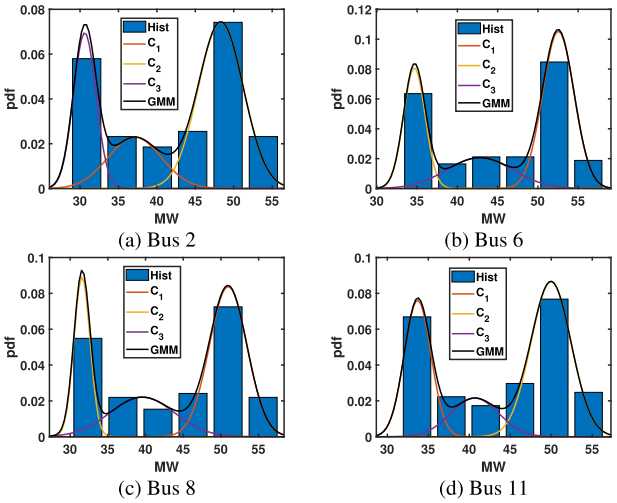


FIGURE 11. 3-Clusters GMM pdf for four load buses.

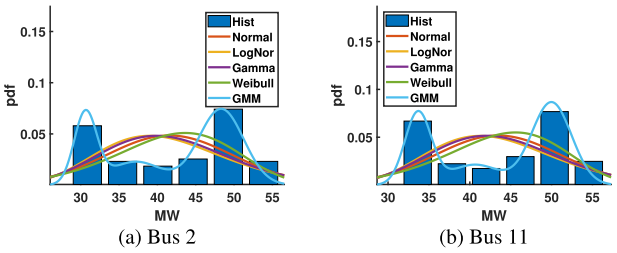


FIGURE 12. Different pdfs, with GMM clusters = 3.

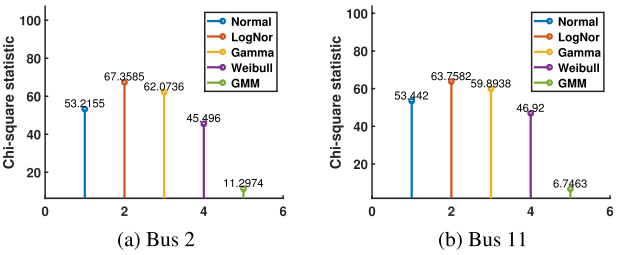


FIGURE 13. Different distributions goodness-of-fit, GMM Clusters = 3.

between the computational efforts of the GMM and accuracy. Commonly GMM contains 3 to 6 clusters because at higher clusters, the value of the Chi-Square statistic is barely changing while the processing time increases incredibly.

One of the benefits of GMM is dividing the loads into groups, i.e., low, medium, and high loads. Each group is represented by at least one Gaussian mixture *GM* (mean and variance), and each *GM* defines the minimum and maximum value of its loads. The GMM variances reflect the uncertainty parameter or weights in the WLS state estimator.

C. STATE ESTIMATION RESULTS AND ACCURACY

The state estimation algorithm is tested on a modified version of the Cigre B4 DC grids test case [40]. The network contains 6 AC systems—22 buses-, 2 DC systems—15 buses-, 8 AC generators (onshore and offshore), 11 converters (rectifiers

TABLE 1. The Cigre B4 test case measurements.

Meas. Type	Count	Details
AC	6 22 & 22 8 & 8	voltages , 1 per AC system active & reactive power injection active & reactive power flow
DC	2 15 4	voltages , 1 per DC system real & zero power injection real power flow
Conv. Power Coupling	11	power constraints, 1 per converter
Conv. Voltage Coupling	11	M_{factor}, 1 per converter

TABLE 2. AC systems: Part of the measurements with bad data.

Meas. Type	True	Noisy	LNR	GMM-LNR
Bus 1: V_{ac}	1.0000	1.0000	1.0000	1.0000
Bus 1: P_{inj}	0.1096	0.1072	0.1080	0.1087
Bus 1: Q_{inj}	−0.0020	−0.0019	0.0022	−0.0022
Bus 2: P_{inj}	−0.1771	−0.1717	−0.1730	−0.1763
Bus 2: Q_{inj}	−0.0115	−0.0114	−0.0119	−0.0116
Bus 5: P_{inj}	0.6420	0.6498	0.6528	0.6427
Bus 5: Q_{inj}	0.2011	0.2101	0.2026	0.2022
Bus 18: P_{inj}	−0.1692	−0.1605	−0.1689	−0.1699
Bus 18: Q_{inj}	0.0000	−1.0000	−0.9985	−0.0012
Bus 5-8: P_{flow}	0.2289	0.2363	0.2338	0.2318
Bus 5-8: Q_{flow}	0.0703	0.0684	0.0701	0.0701
Bus 8-11: P_{flow}	0.0199	0.0198	0.0200	0.0200
Bus 8-11: Q_{flow}	0.0050	0.0053	0.0057	0.0057

and inverters) and 4 demand buses. The network true measurements are obtained by the Julia package PowerModelsACDC [34], [44]. A total 109 measurements were used in the test case as shown in Table 1. The noisy measurements are generated by adding 10% Gaussian noise to all measurements except the slack buses voltages, based on the method in [8]. The LNR threshold is set to be 3, and the number GMM clusters is 3.

The simulations are carried out in three cases:

1) Case I (“NoBD”): a measurement set with 10% Gaussian noise tested in an unified WLS without a pre/post-processing of data.
2) Case II (“BD”): a measurement set with 10% Gaussian error with 2 AC and DC fixed bad data tested in an unified WLS with a LNR post-processing block.
3) Case III (“GMM”): a measurement set with 10% Gaussian error with 2 AC and DC fixed bad data tested in an unified WLS with a GMM pre-processing block and a modified GMM-LNR post-processing block. The RGMM is applied whenever the observed measurement x is connected to more than one Gaussian mixture to ensure that x is connected to a single variance.

Table 2 and Table 3 show part of the AC and DC measurements set for the Case II and III at the first simulation slot (at time 0 of the load profiles). The bad data in the set is indicated by the **bold** buses. In Table 4 the converter measurements set for Case III is presented.

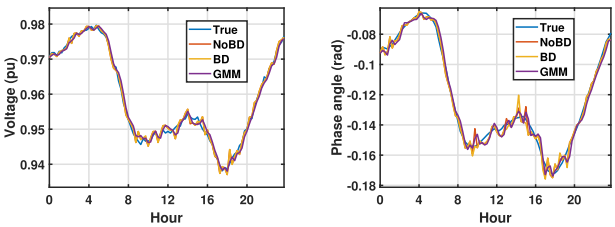
On the AC side, Figs. 14 - 15 show the voltage magnitudes and the phase angle estimations of bus 6 and 11. The results show that the GMM has improved the estimation results despite the 10% noise and the additional bad data. It is noticeable that during high loads (high voltage drop),

TABLE 3. DC systems: Part of the measurements with bad data.

Meas. Type	True	Noisy	LNR	GMM-LNR
Bus 1: V_{dc}	0.9980	0.9980	0.9980	0.9980
Bus 9: V_{dc}	1.0079	1.0079	1.0079	1.0079
Bus 1: P_{inj}	-0.1127	-0.1142	-0.1129	-0.1129
Bus 2: P_{inj}	0.1129	0.1118	0.1131	0.1131
Bus 5: P_{inj}	0.1580	0.1606	0.1559	0.1586
Bus 8: P_{inj}	-0.2371	-0.2277	-0.2329	-0.2336
Bus 11: P_{inj}	-0.0763	0.5000	-0.0755	-0.0754
Bus 15: P_{inj}	0.0818	0.0817	0.0812	0.0812
Bus 3-4: P_{flow}	-0.1518	-0.1476	-0.1494	-0.1497
Bus 6-15: P_{flow}	-0.0096	-0.0095	-0.0071	-0.0090

TABLE 4. GMM-LNR: The converter coupling constraints and measurements.

Meas.	Conv. #	True	Noisy	Estimated
$hpc(x)$	1	0.0	-	-1.73×10^{-5}
	2	0.0	-	-2.01×10^{-6}
	3	0.0	-	-5.84×10^{-5}
	4	0.0	-	-5.42×10^{-5}
	5	0.0	-	-2.08×10^{-4}
	6	0.0	-	-2.16×10^{-6}
	7	0.0	-	-1.73×10^{-5}
	8	0.0	-	-3.13×10^{-5}
	9	0.0	-	1.65×10^{-5}
	10	0.0	-	-5.27×10^{-5}
	11	0.0	-	-6.183×10^{-5}
$hvc(x)$	1	0.734242	0.688934	0.713119
	2	0.705709	0.711326	0.70180
	3	0.714178	0.754003	0.73407
	4	0.73433	0.749302	0.734149
	5	0.714424	0.716197	0.710265
	6	0.730594	0.698595	0.724154
	7	0.707107	0.701619	0.707099
	8	0.716038	0.739066	0.716553
	9	0.715433	0.757917	0.71039
	10	0.716394	0.74327	0.715649
	11	0.733743	0.711731	0.724452



(a) Voltage magnitude: Bus 6 (b) Phase angle: Bus 6

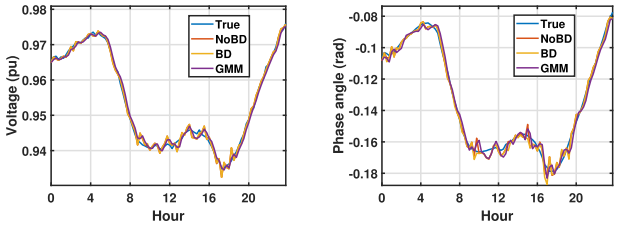
FIGURE 14. Estimated vs true AC voltage magnitude and phase angle.

the estimation gets less accurate since the noise variation is a function of the load magnitude. On the DC side, The GMM has considerably increased the accuracy of the voltage drops estimations as shown in Figs. 16a and 16b for bus 8 and 12.

The accuracy of the state estimation is evaluated using the relative error formula in (29), where $\hat{s_e}$ is the estimated state value and s_t is the actual state value.

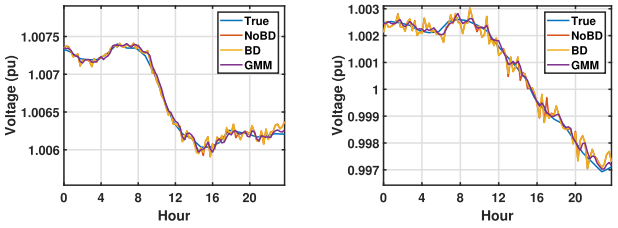
$$\text{RE} = 100\% \times |\frac{\hat{s}_e - s_t}{s_t}| \tag{29}$$

Figs. 17 and 18 show the accuracy of the three estimation cases in some of the AC and DC buses. The figures indicate that the GMM has improved the overall accuracy in both AC



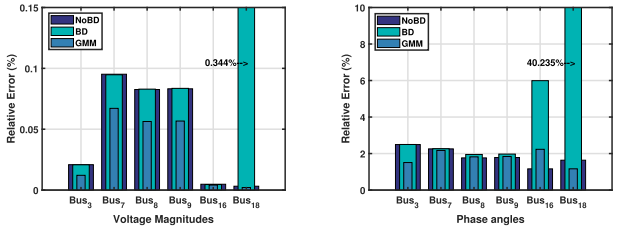
(a) Voltage magnitude: Bus 11 (b) Phase angle: Bus 11

FIGURE 15. Estimated vs true AC voltage magnitude and phase angle.



(a) Voltage magnitude: Bus 8 (b) Voltage magnitude: Bus 12

FIGURE 16. Estimated vs true DC voltage states.



(a) Voltage magnitudes (b) Phase angles

FIGURE 17. State estimation relative error for AC states.

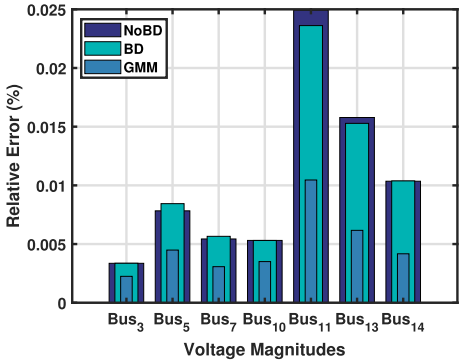


FIGURE 18. State estimation relative error for DC states.

and DC sides. In Fig. 17, the traditional LNR bad data detection failed to accurately correct the measurements around bus 18, causing a considerable diverge in the estimation of the bus states. However, the GMM was able to suspect the bad data around that bus in advance, and afterwards, the GMM-LNR dealt with the measurements.

The state estimation computational efforts are shown in Fig. 19. The average time performance for the three simulation cases are: 0.0313, 0.103 and 0.0761 seconds for Case I, Case II and Case III respectively. The GMM-LNR showed

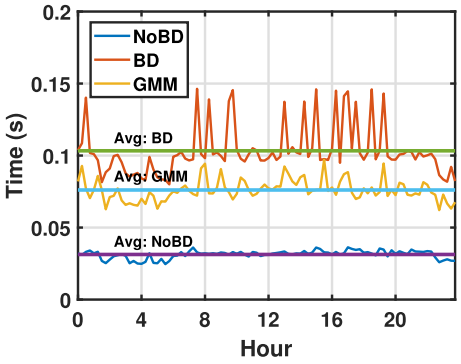


FIGURE 19. The time performance of the three simulation cases.

faster performance than the tradition LNR (a reduction of approx. 26.12%). The figure indicates that between the hour 15 and 19, the GMM-LNR has few computational peaks due to passing some data to the traditional LNR detector as described in Fig. 8. In case II, where LNR is only used, the time performance suffers from higher peaks in the calculations due to the multiple iterations to correct a single bad data.

V. CONCLUSION

The work in this paper has presented a major modification over the traditional LNR bad data detector for WLS algorithm. The modification core is based on GMM to model historical data. The GMM components and parameters were determined using the EM algorithm. The work has integrated the GMM parameters in the unified WLS state estimator for hybrid AC/DC networks. The parameters are used to pre-processing the measurements set to identify the suspected bad data by comparing the set with the modelled GMM mixtures. Afterwards, the state estimator is run and the GMM-LNR is applied on the suspected measurements. Additionally, to achieve higher robustness, the traditional post-processing LNR is applied to any remaining bad measurements, which are not detected in the pre-processing block.

A goodness-of-fit test is used to verify the GMM, and results have proved that the GMM models the load profiles and historical data better than the known distributions.

This work involved three study cases of WLS and measurements set. Each case has shown different results as follows:

- Case I: measurements set was corrupted with 10% Gaussian noise without fixed bad data. The unified WLS was applied to estimate the states. Resulted estimations were reasonably acceptable and fast to converge but fragile to any unexpected bad data.
- Case II: an extension of Case I, whereas measurements included two AC and DC bad data, and WLS robustness was reinforced by the LNR bad data detector block. Results have shown that computational efforts were almost four times the exposed WLS. Also, it provided incorrect estimation when bad data interact and conform with adjacent measurements in the network, as was shown in bus 18 in the AC side.
- Case III: same measurements set of Case II was used. GMM contributions were included in state estimation weights and bad data detection blocks. This approach has outperformed the traditional LNR, results have shown robust estimations against several bad data levels, higher accuracy, and less computational requirements.

REFERENCES

[1] D. Roberson, H. C. Kim, B. Chen, C. Page, R. Nuqui, A. Valdes, R. Macwan, and B. K. Johnson, "Improving grid resilience using high-voltage DC: Strengthening the security of power system stability," *IEEE Power Energy Mag.*, vol. 17, no. 3, pp. 38–47, May 2019.

[2] J. Arrillaga, Y. Liu, and N. Watson, *Flexible Power Transmission: The HVDC Options*. Chichester, U.K.: Wiley, Sep. 2007, pp. 1–362, doi: 10.1002/9780470511862.

[3] Z. Li, R. Zhan, Y. Li, Y. He, J. Hou, X. Zhao, and X.-P. Zhang, "Recent developments in HVDC transmission systems to support renewable energy integration," *Glob. Energy Interconnect.*, vol. 1, no. 5, pp. 595–607, Dec. 2018.

[4] J. Liang, O. Bellmunt, and D. Van Hertem, *HVDC Grid Layouts: For Offshore and Supergrid of the Future*. Hoboken, NJ, USA: Wiley, Mar. 2016, pp. 171–191.

[5] J. W. Lambrechts and S. Sinha, *Microsensing Networks for Sustainable Cities* (Smart Sensors, Measurement and Instrumentation). Denmark: Springer, 2016, pp. 1–339, doi: 10.1007/978-3-319-28358-6.

[6] V. Agelidis, G. Demetriades, and N. Flourentzou, "Recent advances in high-voltage direct-current power transmission systems," in *Proc. IEEE Int. Conf. Ind. Technol.*, Dec. 2006, pp. 206–213.

[7] S. P. Azad, J. A. Taylor, and R. Iravani, "Decentralized supplementary control of multiple LCC-HVDC links," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 572–580, Jan. 2016.

[8] M. Ayiad, H. Leite, and H. Martins, "State estimation for hybrid VSC based HVDC/AC transmission networks," *Energies*, vol. 13, no. 18, p. 4932, Sep. 2020.

[9] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*, vol. 24. New York, NY, USA: Marcel Dekker, Jan. 2004.

[10] J. Zhao, J. Qi, Z. Huang, A. P. S. Meliopoulos, A. Gomez-Exposito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, and A. K. Singh, "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, Jul. 2019.

[11] M. Ingram, S. Bell, S. Mathews, G. J. Cokkinides, and A. P. Meliopoulos, "Use of phasor measurements, SCADA and IED data to improve state estimation procedures," in *Proc. 7th Georgia Tech Fault Disturbance Anal. Conf.*, Jan. 2004, pp. 26–27.

[12] M. B. D. C. Filho, A. M. L. Silva, and D. M. Falcao, "Bibliography on power system state estimation (1968-1989)," *IEEE Trans. Power Syst.*, vol. 5, no. 3, pp. 950–961, Aug. 1990.

[13] C. Carquex, C. Rosenberg, and K. Bhattacharya, "State estimation in power distribution systems based on ensemble Kalman filtering," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6600–6610, Jun. 2018.

[14] A. Sengupta and A. K. Sinha, "A comparative study on state estimation algorithms," in *Proc. Int. Conf. Ind. Inf. Syst.*, Aug. 2007, pp. 317–322.

[15] A. D. L. V. Jaen, E. Acha, and A. G. Exposito, "Voltage source converter modeling for power system state estimation: STATCOM and VSC-HVDC," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1552–1559, Nov. 2008.

[16] E. Unamuno and J. A. Barrena, "Hybrid AC/DC microgrids—Part I: Review and classification of topologies," *Renew. Sustain. Energy Rev.*, vol. 52, pp. 1251–1259, Dec. 2015.

[17] A. Mouco and A. Abur, "A robust state estimator for power systems with HVDC components," in *Proc. North Amer. Power Symp. (NAPS)*, Sep. 2017, pp. 1–5.

[18] T. Zhang, Y. Wang, M. Zhai, Y. Huang, and L. Zhao, "Fast decoupling algorithm of state estimation for hybrid AC/DC power systems with voltage source converter," *Dianli Xitong Zidonghua/Automat. Electr. Power Syst.*, vol. 42, pp. 70–76, Dec. 2018.

[19] X. Yang, H. Zhang, Q. Li, Z. Guo, K. Zhao, X. Li, and F. Han, "A weight modification sequential method for VSC-MTDC power system state estimation," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 207, Jun. 2017, Art. no. 012071.

[20] Q. Ding, T. S. Chung, and B. Zhang, "An improved sequential method for AC/MTDC power system state estimation," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 506–512, Aug. 2001.
[21] E. Zamora, C. F. Esquivel, A. Pizano-Martínez, and H. Estrada-Garcia, "Hybrid state estimator considering SCADA and synchronized phasor measurements in VSC-HVDC transmission links," *Electr. Power Syst. Res.*, vol. 133, pp. 42–50, Apr. 2016.
[22] W. Li, "PMU-based state estimation for hybrid ac and dc grids," Ph.D. dissertation, Electr. Power Energy Syst., KTH, Stockholm, Sweden, 2018.
[23] R. Martinez and C. F. Esquivel, "A new unified approach for the state estimation and bad data analysis of electric power transmission systems with multi-terminal VSC-based HVDC networks," *Electr. Power Syst. Res.*, vol. 160, pp. 251–260, Jul. 2018.
[24] J. Zhao and L. Mili, "Robust unscented Kalman filter for power system dynamic state estimation with unknown noise statistics," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1215–1224, Mar. 2019.
[25] Y. Gu, Z. Yu, R. Diao, and D. Shi, "Doubly-fed deep learning method for bad data identification in linear state estimation," *J. Modern Power Syst. Clean Energy*, vol. 8, no. 6, pp. 1140–1150, 2020.
[26] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, and J. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4766–4778, Nov. 2018.
[27] M. Zhou, Y. Wang, A. K. Srivastava, Y. Wu, and P. Banerjee, "Ensemble-based algorithm for synchrophasor data anomaly detection," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2979–2988, May 2019.
[28] M. Gol and A. Abur, "LAV based robust state estimation for systems measured by PMUs," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1808–1814, Jul. 2014.
[29] X. Kong, Z. Yan, R. Guo, X. Xu, and C. Fang, "Three-stage distributed state estimation for AC-DC hybrid distribution network under mixed measurement environment," *IEEE Access*, vol. 6, pp. 39027–39036, 2018.
[30] A. Foroutan and F. Salmasi, "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.*, vol. 2, no. 4, pp. 161–171, Jul. 2017.
[31] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, Oct. 2010.
[32] A. K. Singh, S. K. Ibraheem, and M. Muazzam, "An overview of electricity demand forecasting techniques," in *Proc. Nat. Conf. Emerg. Trends Elect., Instrum. Commun. Eng.*, vol. 3, no. 3, 2013, pp. 38–48.
[33] M. Ayiad, H. Martins, O. Nduka, and B. Pal, "State estimation of low voltage distribution network with integrated customer-owned PV and storage unit," in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–6.
[34] H. Ergun, J. Dave, D. Van Hertem, and F. Geth, "Optimal power flow for AC–DC grids: Formulation, convex relaxation, linear approximation, and implementation," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2980–2990, Jul. 2019.
[35] E. Acha, P. Roncero-Sánchez, A. de la Villa-Jaen, L. Castro, and B. Kazemtabrizi, *VSC-FACTS-HVDC: Analysis, Modelling and Simulation in Power Grids*. Hoboken, NJ, USA: Wiley, Apr. 2019.
[36] E. Handschin, F. C. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Trans. Power App. Syst.*, vol. APS-94, no. 2, pp. 329–337, Mar. 1975.
[37] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Germany: Springer-Verlag, 2006.
[38] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
[39] D. J. Salmond, "Mixture reduction algorithms for target tracking in clutter," *Proc. SPIE*, vol. 1305, p. 434, Oct. 1990, doi: 10.1117/12.2321784.
[40] T. Vrana, S. Dennetiere, Y. Yang, J. A. Jardini, D. Jovcic, and H. Saad, "The Cigré B4 DC grid test system," *CIGRE Electra*, vol. 270, pp. 1–12, Oct. 2013.
[41] NationalGridESO. (Jan. 2021). *National Grid ESO: Historic Demand Data*. [Online]. Available: https://data.nationalgrideso.com/demand/historic-demand-data
[42] D. W. Scott, "Scott's rule," *Wiley Interdiscipl. Rev., Comput. Statist.*, vol. 2, no. 4, pp. 497–502, Jul. 2010. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.103
[43] G. A. P. Cirrone, S. Donadio, S. Guatelli, A. Mantero, B. Mascialino, S. Parlati, M. G. Pia, A. Pfeiffer, A. Ribon, and P. Viarengo, "A goodness-of-fit statistical toolkit," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 5, pp. 2056–2063, Oct. 2004.
[44] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM Rev.*, vol. 59, no. 1, pp. 65–98, Jan. 2017, doi: 10.1137/141000671.

MOTAZ AYIAD received the B.S. degree in mechatronics engineering from Al-Azhar University-Gaza (AUG), Palestine, in 2015, and the M.S. degree in future power networks from Imperial College London, in 2017. In 2015/2016, he worked as a Teaching Assistant at AUG and as a Smartmeter Engineer at WebCom Technology. Since 2018, he has been a member of the EU-funded InnoDC Project, a Marie Curie Ph.D. Researcher with the University of Porto, and a Researcher at Efacec, Maia, Portugal. During his journey in the InnoDC Project, he participated in several policies works, public events, and tutorials.

HELDER LEITE received the degree in electrical engineering from the University of Porto, Portugal, in 2000, and the Ph.D. degree in electrical engineering from the University of Manchester, Manchester, U.K., in 2004. He has been a Lecturer with the University of Porto, since 2005. His research focus includes in the areas of power systems, more specifically power system protection and automation (such as IEC 61850, smart grid, protection testing, automation, and control), asset management and condition monitoring, distributed energy resources integration, and VSC-based HVDC projects.

HUGO MARTINS graduated in electrical and computer engineering from the University of Porto, in 1998. He joined Efacec, in 1999. He currently works as a SCADA Product Manager with the Network Management Division, part of the Automation and Energy Systems Business Unit. Before 2017, he worked as a Software Engineer in the specification and development of ScateX# (SCADA/ADMS network management system). Since 2013, he has been participated in several funded projects, both national (P2020) and international (H2020), as a Technical Manager and Coordinator for Efacec.