IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 April 2023, accepted 17 April 2023, date of publication 26 April 2023, date of current version 4 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3270386

RESEARCH ARTICLE

Robust State Estimation Model for Low Voltage Distribution Networks in the Presence of Multiple Gross Errors

YASSINE BOUKILI[1], MOTAZ AYIAD[2], HAMED MOAYYED[3], A. PEDRO AGUIAR[1], (Senior Member, IEEE), AND ZITA VALE[3], (Senior Member, IEEE)

[1]SYSTEC-ARISE, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal
[2]PAC Research and Development Division, Efacec Automation, 4471-907 Porto, Portugal
[3]GECAD-LASI, Polytechnic of Porto, 4249-015 Porto, Portugal

Corresponding author: Yassine Boukili (up201711573@fe.up.pt)

This work was supported in part by the base funding (UIDB/00147/2020) and programmatic funding (UIDP/00147/2020) of the SYSTEC—Center for Systems and Technologies; in part by ARISE—Associate Laboratory for Advanced Production and Intelligent Systems (LA/P/0112/2020) and the RELIABLE project (PTDC/EEI-AUT/3522/2020), both funded by national funds through the Fundação para a Ciência e a Tecnologia (FCT)/Ministério da Ciência, Tecnologia e Ensino Superior (MCTES) Programa de Investimentos e Despesas de Desenvolvimento da Administração Central (PIDDAC); in part by the Fundo Europeu de Desenvolvimento Regional (FEDER) funds through COMPETE program and the national funds through FCT under Project UIDP/00760/2020; and in part by the Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD) Research Center (UIDB/00760/2020) for providing work facilities and equipment to the project team.

ABSTRACT One of the primary challenges in solving the State Estimation (SE) problem in low voltage networks is the presence of Gross Errors (GE). If the SE model fails to accurately estimate state variables in the presence of GE, the system operator may receive a distorted image of the power network, potentially leading to unforeseen disruptions, blackouts, and significant economic losses. The classical weighted least squares method, which is commonly used in such problems, exhibits low accuracy when simultaneous GE occurs. This paper provides a comprehensive analysis comparing robust M-estimators designed to handle GE, such as Hachtel and the largest normalized residual test, and presents a novel SE method called the Adaptive Maximum Correntropy Criterion (AMCC). The AMCC employs the maximum correntropy criterion for the loss function and an interior point methodology. Additionally, an adaptive scheme is employed to automatically adjust a high parameter associated with the shape of the related gamma function. We show that, in a real low voltage network, the AMCC exhibits superior accuracy with smaller root-mean-square errors compared to the other estimators studied.

INDEX TERMS Cauchy, Hachtel, interior point method, largest normalized residual test, least square, least absolute value, low voltage, maximum correntropy criterion, M-estimator, non-Gaussian noises, state estimation.

I. INTRODUCTION

The State Estimation (SE) is a process in power systems networks that aims to provide accurate system states, such as the voltage angle θ and amplitude V [1]. SE uses measurements acquired from the network through various mechanisms, including Phasor Measurement Unit (PMU) devices, being one of the most accurate sources of measurements, followed by intelligent electronic devices, supervisory control and data acquisition systems, and pseudo-measurements generated from historical data being this one the least accurate [2]. Recently, distribution networks have been bolstered by the installation of smart meter infrastructure, resulting in increased network observability and more reliable measurements [5]. Smart meters can store data for a specific time period [6], and provide superior observability in comparison to pseudo measurements. They also improve the accuracy of network modeling and measurements, as well as timestamp synchronization and data acquisition through modern, fast communication networks. Smart meters typically have

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

an accuracy of 10%, whereas PMUs have an accuracy of 1% [7]. In certain distribution networks, the accuracy of state estimation is influenced by the network topology; for instance, radial distribution networks face a greater observability burden than mesh networks [8]. However, these measurements are subject to various factors that can contaminate or invalidate their accuracy, such as temperature changes, sensor anomalies, and sudden system modifications. These factors introduce Gaussian and/or non-Gaussian noise in the collected data [3], [4].

The classical Weighted Least Squares (WLS) method is widely used in power systems due to its reliability, particularly when the network is fully observable [9]. However, the WLS method is sensitive to the presence of Gross Errors (GE). When GE contaminates the data, two approaches can be applied: posterior and prior. The posterior approach [14], also known as Largest Normalized Residual Test (LNRT) [13], is a classical GE detection tool associated with the WLS method. It diagnoses the residual (the difference between estimated and true values) to identify contaminated measurements with GE, eliminating or correcting them while maintaining network observability. The LNRT approach requires a re-run of the SE with the updated measurement set. In some cases, pseudo-measurements from historical data are used to substitute the eliminated measurements and maintain observability. The prior approach [15] involves applying SE directly to the contaminated data, thereby correcting it without requiring further iterations as in the previous approach.

In [16], a method is proposed that considers a reduced number of measurements and low-accuracy pseudo-measurements to increase system observability. In smart grids, a data-driven power flow method is proposed based on a learning model that utilizes historical or simulated data, constructed to remove the effect of data collinearity [17]. Another data-driven robust state estimation method is proposed in [18] to address the shortcomings of model-driven approaches.

A variety of SE methods have been proposed to address outliers. For example, [19] uses a linearized error analysis to identify GE, while [2], [20] propose innovative approaches for GE detection, identification, and correction. A new method based on a generalized pattern search algorithm for data correction in harmonic SE is presented in [21]. In the SE field, M-estimators have also been proposed in the literature, such as $L1$ loss (often cited as Charbonnier loss in [22]), pseudo-Huber loss [23], $L1$-$L2$ loss [24], German-McClure [25], Cauchy method [26], and Welsch loss [27].

In [28], a maximum exponential absolute value SE model is proposed that uses Laplace kernel and the Maximum Correntropy Criterion (MCC) in the objective function to improve robustness. A Primal-Dual Interior Point Methodology (IPM) is employed to enhance accuracy. Recent works such as [29] and [30] also use a Generalized Correntropy Criteria (GCC) to improve robustness and control the induced metric's behavior. These approaches offer flexible parameters, unlike traditional M-estimators such as $L1$ and $L2$ loss functions. In [31], a MCC SE model is proposed that is based on the GCC and does not require explicit GE detection. The GCC-based methods are further improved in [32], [33], and [34] using a Generalized Correntropy Interior Point (GCIP) method. However, manually setting the kernel shape and standard deviation parameters is not practical in real-world applications.

A. CONTRIBUTION

Overall, when there are enough measurements, the WLS method offers precise estimations, producing results that closely align with the true values. However, it can be highly sensitive to GE, which can cause the state estimate to exceed reasonable boundaries.

This paper proposes a new robust state estimation (SE) method called the Adaptive Maximum Correntropy Criterion (AMCC), which is an improvement over the GCIP method [34]. The proposed AMCC utilizes a flexible Generalized Correntropy Induced Metric (GCIM) together with an effective scheme that provides the ability to adapt to various datasets in successive iterations. Unlike the GCIP, which requires manual setting of the shape parameter "c" for each static snapshot, the AMCC introduces an adaptive scheme to automatically adjust the parameter associated with the shape of the related gamma function. This allows for successive application to static snapshots, making it more practical for use. A comparative study of performance is conducted between various M-estimators that address outliers, including Cauchy, Weighted Least Absolute Value (WLAV), Hachtel, WLS with LNRT, and GCIP. In this work, the LNRT has been fine-tuned to reduce the occurrence of GEs in the data. This was achieved by selectively removing suspected measurements, while retaining others, resulting in high accuracy with minimal processing time. However, it is important to note that implementing the LNRT involves using multiple auto-encoders that need to be well-tuned and retrained if there are any changes in the network topology [32]. Moreover, the LNRT method may not be effective in cases where successive GEs are deleted since this can result in a loss of network observability [43], the issue described does not occur when applying the GCIP and AMCC methods.

To compare the estimation performance, a real Low Voltage (LV) load data profile in Kent, England [36] is utilized. The true measurements of this network are obtained from power flow calculations using actual load profiles. The measurement sets for the case studies are generated by corrupting the true measurements with Gaussian noise and GE. Three data scenarios with different percentages of GE in the total measurement data set are examined: 4%, 8%, and 16%. The mean residuals for each of the studied SE methods are then calculated. To address the economic aspect of SE (by using fewer meters), the simulation uses a minimum number of measurements. The results demonstrate that the AMCC outperforms other SE methods in terms of estimation accuracy, exhibiting smaller Root-Mean-Square Errors (RMSE).

The paper is structured as follows: Section II sets out the problem and presents the studied M-estimators, including the WLS, LNRT, Cauchy, WLAV, and Hachtel methods constrained by the zero-injection function. In Section III, correntropy-based SE methods are presented, with subsections A and B sequentially discussing the GCIP method and the AMCC model. Section IV is devoted to the case study, where the network description, load profile data, study data cases, and simulation results are presented and discussed.

II. ROBUST M-ESTIMATORS

The Maximum Likelihood (ML) formulation was one of the first models for data regression in SE, introduced by Fisher between 1912 and 1922 [37]. It was based on ideas discussed by Daniel Bernoulli [38] and Gauss [39]. By applying a logarithm operator, the maximization problem of the ML formulation can be expressed in various M-estimator forms, as presented below.

The observed measurements consist of an m-dimensional vector $z = [z_1\ z_2\ \dots\ z_m]$, which includes the line power flows, bus power injections, bus voltage magnitudes, and line current flow magnitudes. The residual is therefore an m-dimensional measurement error vector, as defined in [40],

$$r(x) = h(x) - z \tag{1}$$

with $x \in \mathbb{R}^s$, where s is the number of state variables, namely θ and V, for all buses. The measuring functions constituting $h(x)$ are constructed using active/reactive power injection and flow, current flow, and voltage magnitude, as described in [41].

A. WLS METHOD WITH LNRT

1) THE WLS METHOD

In this classical optimization problem, the objective is to find a solution that minimizes the sum of the weighted squares of the residuals. Here, a residual is defined as the difference between the predicted value provided by the model and the observed measurement. Assuming that the error has a zero mean and covariance R, the classical WLS SE method is solved by minimizing the weighted least squares criterion, with $W = R^{-1}$, as described in [1],

$$\min_x J_1(x) = \frac{1}{2}\, r(x)^T W r(x) \tag{2}$$

In [42], the zero-injection constraint $g(x) \in \mathbb{R}^n$, where n is the number of equality constraints, is associated with the WLS objective function to address the issue of outliers. In this scenario, we have (for diagonal R):

$$\begin{cases} \min\limits_x J_1(x) = \frac{1}{2} \sum\limits_{i=1}^{m} W_{ii}\, r_i(x)^2 \\ s.t.\ \ g(x) = 0 \end{cases} \tag{3}$$

2) THE LNRT

This is a conventional method for bad data detection that is combined with the WLS method to check for the presence of gross errors in the WLS outputs. This method eliminates or corrects a single gross error at each iteration and is typically accurate when there are minimal gross errors in the data. However, it may lead to a loss of observability due to the removal of measurements. The steps involved in the LNRT procedure are presented in Table 1.

TABLE 1. LNRT method.

Step 1: The presence of gross errors in the data can be verified using the Chi-square test χ^2, as described in [12];
Step 2: The LNRT parameter (r^{GE}) can be calculated for all measurements using the following procedure:

$$r^{GE} = \sum_{i=1}^{m} \frac{|\, r_i \,|}{\sqrt{S_{ii}}} \tag{4}$$

where

$$S = R - \left(H \left(H^T R^{-1} H \right)^{-1} H^T \right), \quad H = \sum_{i=1}^{m} \frac{\partial h_i(x)}{\partial x} \tag{5}$$

Step 3: For a certain threshold (t) usually between 1.5 and 4 (3 in our simulation case) that depends on the sensibility of the network, the suspected measurements are deleted from the data, where $r^{GE} > t$, as long as the network observability is guaranteed [43];
Step 4: Restart the process from the WLS algorithm with updated measurements;

B. OBJECTIVE FUNCTIONS

1) CAUCHY METHOD

The following estimator is derived from the Cauchy distribution, as described in [44]

$$\begin{cases} \min\limits_x J_2(x) = \sum\limits_{i=1}^{m} ln(1 + r_i(x)^2) \\ s.t.\ \ g(x) = 0 \end{cases} \tag{6}$$

2) WLAV METHOD

The L1 or double exponential estimator, which is based on the Laplacian distribution, is also known as described in [45]

$$\begin{cases} \min\limits_x J_3(x) = \sum\limits_{i=1}^{m} W_{ii}\, |r_i(x)| \\ s.t.\ \ r(x) = h(x) - z \\ g(x) = 0 \end{cases} \tag{7}$$

3) HACHTEL METHOD

The optimization problem of Hachtel can be expressed as given in [46]

$$\begin{cases} \min\limits_x J_4(x) = \frac{1}{2} \sum\limits_{i=1}^{m} W_{ii}\, r_i(x)^2 \\ s.t.\ \ r(x) = h(x) - z \\ g(x) = 0 \end{cases} \tag{8}$$

C. FORMULATION OF THE M-ESTIMATORS

1) NEWTON METHOD

A Newton iterative procedure can be developed for the related Lagrangian function to solve the WLS and Cauchy objective

functions. In this case,

$$L(x) = \sum_{i=1}^{m} J_i(x) - \sum_{j=1}^{n} \beta_j \, g_j(x) \tag{9}$$

where $\beta \in \mathbb{R}^n$ is a Lagrange multiplier. After computing the gradient (∇L) and Hessian ($\nabla^2 L$) of the Lagrangian, the following equation is computed to determine the steps $dx \in \mathbb{R}^s$ and $d\beta \in \mathbb{R}^n$ [47]

$$\begin{bmatrix} \nabla^2_{xx} L & \nabla^2_{x\beta} L \\ \nabla^2_{\beta x} L & \nabla^2_{\beta\beta} L \end{bmatrix} \begin{bmatrix} dx \\ d\beta \end{bmatrix} = \begin{bmatrix} -\nabla_x L \\ -\nabla_\beta L \end{bmatrix} \tag{10}$$

where the new iteration $k+1$ is deduced,

$$\begin{bmatrix} x_{k+1} \\ \beta_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \beta_k \end{bmatrix} + \begin{bmatrix} dx \\ d\beta \end{bmatrix} \tag{11}$$

In Hachtel method case, the adopted Lagrangian function is written as

$$\begin{aligned} \mathcal{L}_4 \equiv & \sum_{i=1}^{m} J_{4i}(x) - \sum_{j=1}^{n} \beta_j g_j(x) \\ & - \sum_{i=1}^{m} \alpha_i (r_i - h_i(x) + z_i) \end{aligned} \tag{12}$$

where $\alpha \in \mathbb{R}^m$.

As described in [48], once the gradients of the Lagrangian with respect to x, β, α, and r are computed, the following can be deduced:

$$\begin{bmatrix} -G & 0 & 0 \\ 0 & -G^T & H^T \\ H & 0 & -W^{-1} \end{bmatrix} \begin{bmatrix} dx \\ d\beta \\ d\alpha \end{bmatrix} = \begin{bmatrix} g(x) \\ G^T\beta - H^T\alpha \\ W^{-1}\alpha + z - h \end{bmatrix} \tag{13}$$

where $d\alpha \in \mathbb{R}^m$ and $G = \sum_{j=1}^{n} \frac{\partial g_j(x)}{\partial x}$. The new iteration is computed following Newton method as

$$\begin{bmatrix} x_{k+1} \\ \beta_{k+1} \\ \alpha_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \beta_k \\ \alpha_k \end{bmatrix} + \begin{bmatrix} dx \\ d\beta \\ d\alpha \end{bmatrix} \tag{14}$$

2) INTERIOR POINT METHOD

The presence of absolute values in residuals complicates standard optimization procedures such as the Newton method, which necessitates the use of Linear Programming (LP) methods. Two typical LP strategies, namely simplex and IPM, have been applied to solve optimization problems so far. The difference between the two strategies lies in the way of reaching the solution: in the simplex strategy, the extreme point is traced along the exterior of the feasible region, while in the IPM, it is traced through an interior path [1]. In both strategies, slack variables could be used to reformulate the optimization problem and eliminate the absolute value.

Let Υ be defined such that: $|r| < \Upsilon$, it could be replaced by two equalities when introducing two non-negative slack variables $l, k \in \mathbb{R}^m$ [1]:

$$r_i - l_i = -\Upsilon_i \tag{15}$$

$$r_i + k_i = \Upsilon_i \tag{16}$$

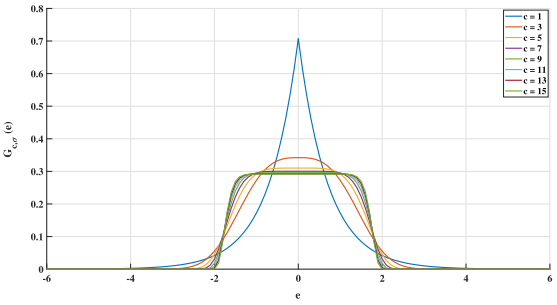


FIGURE 1. GGD with different values of c and $\varsigma = 1$.

then, one can deduce the following equations:

$$r_i = u_i - v_i \tag{17}$$

$$\Upsilon_i = u_i + v_i \tag{18}$$

where $u_i = \frac{k_i}{2}$ and $v_i = \frac{l_i}{2}$. Hence, the following modification will be taken the optimization problem of WLAV method

$$\begin{cases} \min\limits_{u,v} J_3(u, v) = \sum\limits_{i=1}^{m} W_{ii} \, (u_i + v_i) \\ s.t. \;\; r(x) - u + v = 0 \\ g(x) = 0 \end{cases} \tag{19}$$

The adopted formulation of this objective function has been presented and discussed in [1].

III. THE ADAPTIVE MAXIMUM CORRENTROPY CRETERION

A. GENERALIZED CORRENTROPY

The correntropy can be defined as a generalized measure of similarity between two Probability Density Functions (PDF), as noted in [30]. It has a strong relationship with entropy, as described in [49] and [50].

Consider two scalar random variables X and Y, the correntropy equation can be expressed by

$$\begin{aligned} \upsilon_\omega(X, Y) &= E[\kappa_\omega(X - Y)] \\ &= \iint_{xy} \kappa_\omega(X - Y) \, p(x, y) \, \mathrm{d}x\mathrm{d}y \end{aligned} \tag{20}$$

where κ_ω is a kernel operator with parameter (width) ω.

The correntropy is based on the Gaussian Kernel, symmetric, bounded and positive. In case it is based on the Generalized Gaussian Distribution (GGD), is called the GCC [29]. The distribution of the GGD is writen as

$$G_{c,\omega}(r) = \frac{c}{2\,\omega\,\Gamma(1/c)} \exp\left(-\left|\frac{r}{\omega}\right|^c\right) \tag{21}$$

where $\Gamma(.)$ is the gamma function, $c > 0$ is the shape parameter, r is the error between X and Y (the residual), ω is the scale parameter (bandwidth), and it can be expressed by

$$\omega = \varsigma \sqrt{\frac{\Gamma(1/c)}{\Gamma(3/c)}} \tag{22}$$

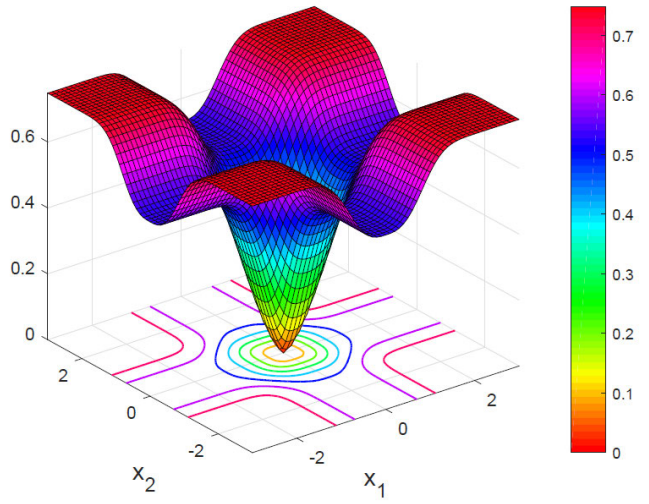


FIGURE 2. GCIM surface in 3*D* space when $c = 3$ and $\omega = 1$.

where ς is a standard deviation of the kernel. The GGD curve is presented in Fig. 1 for different values of c while $\varsigma = 1$. Note that when $c = 1$ we have a Laplace density function, with $c = 2$ a Gaussian function, and when $c \to +\infty$ the uniform distribution over $(-\omega, \omega)$.

Given a finite number of samples S, the GCC can be estimated using the sample estimator

$$\hat{v}_{c,\omega}(X, Y) = \frac{1}{S} \sum_{i=1}^{S} G_{c,\omega}(x_i - y_i) \tag{23}$$

The GCIM, which is related to the GCC, exhibits various L metrics for measuring distance in space, depending on the chosen parameters and the analysis region. This flexibility is utilized to distinguish outliers. Figure 2 demonstrates the varying shapes of the metric in different regions by plotting the distance from the origin using the GCIM measure in a $2D$ space, where $c = 3$ and $\omega = 1$. Note that these parameter values are provided solely for illustrative purposes, and during the auto-tuning process, GCIM can take on other forms to ensure flexibility and accuracy.

1) OBJECTIVE FUNCTION

We now formulate the SE method by casted as the following optimization problem

$$\begin{cases} \max\limits_{x} F(x) = \frac{1}{m} \sum\limits_{i=1}^{m} \dfrac{c}{2\,\omega_i\,\Gamma(1/c)} \exp\left(-\left|\dfrac{r_i(x)}{\omega_i}\right|^c\right) \\ s.t.\ g(x) = 0 \end{cases} \tag{24}$$

In order to deal with the absolute value, we introduce positive slack variables p and q along with a pair of inequality constraints. This yields the following expression:

$$\begin{cases} \max\limits_{p,q} F(p, q) = \frac{1}{m} \sum\limits_{i=1}^{m} \dfrac{c}{2\,\omega_i\,\Gamma(1/c)} \exp\left(-\left(\dfrac{p_i + q_i}{\omega_i}\right)^c\right) \\ s.t.\ f(x, p, q) = r(x) - p + q = 0 \\ \quad g(x) = 0 \\ \quad p, q > 0 \end{cases} \tag{25}$$

where x is a feasible point if it satisfies all the constraints in (25). The feasible region is the set built from all the feasible points.

2) GENERAL FORMULATION OF GCC

The following Lagrangian function can be deduced from the (25)

$$\begin{aligned} \mathcal{L} \equiv\ & F(p, q) + \sum_{j=1}^{n} \beta_j g_j(x) \\ & + \sum_{i=1}^{m} (\alpha_i f_i(x, p, q) + \lambda_i\, p_i + \gamma_i\, q_i) \end{aligned} \tag{26}$$

where $\beta \in \mathbb{R}^n$ and $\alpha,\ \lambda,\ \gamma \in \mathbb{R}^m$, denote the Lagrange multipliers, where the values of all variables will be discussed in the following formulation (overall algorithm). Additional terms are associated to the problem, taken from the Karush-Kuhn-Tucker (KKT) optimality conditions [50]

$$\mathcal{L}_{\lambda_i} \equiv \lambda_i\, p_i = 0 \tag{27}$$

$$\mathcal{L}_{\gamma_i} \equiv \gamma_i\, q_i = 0 \tag{28}$$

To incorporate relaxation, we introduce a perturbation parameter $\mu > 0$, where μ is defined as $\mu = \frac{\rho\,\Delta}{2\,m}$. Here, Δ is the duality gap, which is given by $\Delta = \gamma^T q + \lambda^T p$, and $\rho \in (0, 1)$ is the centering parameter as proposed by [51]

$$\mathcal{L}^{\mu}_{\lambda_i} \equiv \lambda_i\, p_i - \mu = 0 \tag{29}$$

$$\mathcal{L}^{\mu}_{\gamma_i} \equiv \gamma_i\, q_i - \mu = 0 \tag{30}$$

The gradients deduced from the Lagrangian function can be written as

$$\nabla_x \mathcal{L} = G\,\beta + H\,\alpha \tag{31}$$

$$\nabla_{\alpha_i} \mathcal{L} = r_i(x) - p_i + q_i \tag{32}$$

$$\nabla_{\beta_j} \mathcal{L} = g_j(x) \tag{33}$$

$$\begin{aligned} \nabla_{p_i} \mathcal{L} \equiv\ & F(p, q) \left(\frac{-c(p_i + q_i)^{c-1}}{m\,\omega_i^c}\right) \exp\left(-\left(\frac{p_i + q_i}{\omega_i}\right)^c\right) \\ & - \alpha_i + \lambda_i = 0 \end{aligned} \tag{34}$$

$$\begin{aligned} \nabla_{q_i} \mathcal{L} \equiv\ & F(p, q) \left(\frac{-c(p_i + q_i)^{c-1}}{m\,\omega_i^c}\right) \exp\left(-\left(\frac{p_i + q_i}{\omega_i}\right)^c\right) \\ & + \alpha_i + \gamma_i = 0 \end{aligned} \tag{35}$$

By combining the Newton method with the KKT conditions obtained, we arrive at the following result:

$$-\mathcal{L}^{\mu}_{\lambda_i} = \lambda_i\, \mathrm{d}p_i + p_i\, \mathrm{d}\lambda_i \tag{36}$$

$$-\mathcal{L}^{\mu}_{\gamma_i} = \gamma_i\, \mathrm{d}q_i + q_i\, \mathrm{d}\gamma_i \tag{37}$$

The hessian in Newton method yields

$$-\nabla_{p_i}\mathcal{L} = e_i\,\mathrm{d}p_i + e_i\,\mathrm{d}q_i - \mathrm{d}\alpha_i + \mathrm{d}\lambda_i \tag{38}$$
$$-\nabla_{q_i}\mathcal{L} = e_i\,\mathrm{d}p_i + e_i\,\mathrm{d}q_i + \mathrm{d}\alpha_i + \mathrm{d}\gamma_i \tag{39}$$
$$-\nabla_{\beta_j}\mathcal{L} = G_j\,\mathrm{d}x_j \tag{40}$$
$$-\nabla_{\alpha_i}\mathcal{L} = H_i\,\mathrm{d}x_i - \mathrm{d}p_i + \mathrm{d}q_i \tag{41}$$
$$-\nabla_x\mathcal{L} = G_j\,d\beta_j + H_i\,d\alpha_i \tag{42}$$

where the hessian matrices of $h(x)$ and $g(x)$ $\left(\nabla^2 h(x) \text{ and } \nabla^2 g(x)\right)$ are neglected in the (42) and

$$e_i = \frac{F(p,q)}{m}\left[\frac{-c(c-1)(p_i+q_i)^{c-2}}{\omega^c}\exp\left(-\left(\frac{p_i+q_i}{\omega}\right)^c\right) + \frac{c^2(p_i+q_i)^{2c-2}}{\omega^{2c}}\exp\left(-\left(\frac{p_i+q_i}{\omega}\right)^c\right)\right] \tag{43}$$

Substituting (38) and (39) into (36) and (37), the following is deduced

$$\begin{aligned} &-p_i\,e_i\,dq_i + (-p_i\,e_i + \lambda_i)\,dp_i \\ &= -p_i\,d\alpha_i + p_i\,\nabla_{p_i}\mathcal{L} - \mathcal{L}^{\mu}_{\lambda_i} \end{aligned} \tag{44}$$
$$\begin{aligned} &(-q_i\,e_i + \gamma_i)\,dq_i - q_i\,e_i\,dp_i \\ &= q_i\,d\alpha_i + q_i\,\nabla_{q_i}\mathcal{L} - \mathcal{L}^{\mu}_{\gamma_i} \end{aligned} \tag{45}$$

Let a, b, c and $d \in \mathbb{R}^m$, be selected such that

$$\begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} = \begin{bmatrix} -p_i\,e_i & -p_i\,e_i + \lambda_i \\ -q_i\,e_i + \gamma_i & -q_i\,e_i \end{bmatrix}^{-1} \tag{46}$$

Then from (44) and (45) we obtain

$$\begin{bmatrix} dq_i \\ dp_i \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}\begin{bmatrix} -p_i\,d\alpha_i + p_i\ \nabla_{p_i}\mathcal{L} - \mathcal{L}^{\mu}_{\lambda_i} \\ q_i\,d\alpha_i + q_i\ \nabla_{q_i}\mathcal{L} - \mathcal{L}^{\mu}_{\gamma_i} \end{bmatrix} \tag{47}$$

It follows that

$$\mathrm{d}q_i = n_{1i}\,\mathrm{d}\alpha_i + t_{1i} \tag{48}$$
$$\mathrm{d}p_i = n_{2i}\,\mathrm{d}\alpha_i + t_{2i} \tag{49}$$

where

$$n_{1i} = -a_i\,p_i + b_i\,q_i \tag{50}$$
$$n_{2i} = -c_i\,p_i + d_i\,q_i \tag{51}$$
$$t_{1i} = a_i\,(p_i\,\nabla_{p_i}\mathcal{L} - \mathcal{L}^{\mu}_{\lambda_i}) + b_i\,(q_i\,\nabla_{q_i}\mathcal{L} - \mathcal{L}^{\mu}_{\gamma_i}) \tag{52}$$
$$t_{2i} = c_i\,(p_i\,\nabla_{p_i}\mathcal{L} - \mathcal{L}^{\mu}_{\lambda_i}) + d_i\,(q_i\,\nabla_{q_i}\mathcal{L} - \mathcal{L}^{\mu}_{\gamma_i}) \tag{53}$$

By substituting (48) and (49) into (41), one can obtain

$$H\,\mathrm{d}x + C\,\mathrm{d}\alpha = \nu \tag{54}$$

where $C \in \mathbb{R}^{m\times m}$ is a diagonal matrix and

$$C_i = -n_{1i} + n_{2i} \tag{55}$$
$$\nu = z - h(x) + p - q + t_1 - t_2, \quad t_1, t_2 \in \mathbb{R}^m \tag{56}$$

TABLE 2. AMCC algorithm.

GGD parameters initialization $c_{ini} = 1$, $\varsigma = 0.995$ and $flag = 0$;
Admittance vector construction: Y (Tab. 3);
while $flag = 0$
Bandwidth parameters calculation: ω (Equ. 22);
Lagrangian multiplier vectors initialization: $\alpha = \gamma = \lambda = \mathbf{1}_m$ and $\beta = \mathbf{1}_n$;
Non-negative slack variables initialization: $p = q = \mathbf{1}_m$;
Gap initialization: $\Delta = \sum_{i=1}^{m} \gamma_i\,q_i + \lambda_i\,p_i$;
while $\Delta > 10^{-8}$
Initialization of the state variables;
Perturbation factor calculation: $\mu = \frac{\rho\Delta}{2m}$;
Jacobian calculation: $H(x) = [{}^{\partial h(x)}/_{\partial\theta},\ {}^{\partial h(x)}/_{\partial V}]$;
Calculation of: $G(x) = [{}^{\partial g(x)}/_{\partial\theta}, {}^{\partial g(x)}/_{\partial V}]$, where $g(x)$ is the zero-injection function;
Generalized correntropy parameters calculation: e_i (43); a_i, b_i, c_i and d_i (46); n_{1i} (50) and n_{2i} (51); $\mathcal{L}^{\nu}_{\lambda_i}$ (29) and $\mathcal{L}^{\nu}_{\gamma_i}$ (30); $\nabla_{p_i}\mathcal{L}$ (34); $\nabla_{q_i}\mathcal{L}$ (35); t_{1i} (52) and t_{2i} (53);
Calculation of dx, $d\beta$ and $d\alpha$ vectors (57);
Calculation of dq_i (48) and dp_i (49); $d\gamma_i$ (39) and $d\lambda_i$ (38);
Primal and dual step-length calculation: Δ_P and Δ_Q;
Updating the values of p_i, q_i, α_i, β_i, γ_i, λ_i (58) and Δ.
end while
Substituting ν (56) into the residual $r(x)$ (59). For multiple c values, c_{new} is deduced when ${}^{\partial F}/_{\partial c} = 0$;
if $c_{ini} = c_{new}$
$flag = 1$;
else
$c_{ini} = c_{new}$;
end if
end while
return x

If we use equations (33) and (40), along with (31) and (42), and also (54), we can figure out the following linear equation:

$$\begin{bmatrix} G & 0 & 0 \\ 0 & G^T & H^T \\ H & 0 & C \end{bmatrix}\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}\beta \\ \mathrm{d}\alpha \end{bmatrix} = \begin{bmatrix} -\nabla_{\beta}\mathcal{L} \\ -\nabla_x\mathcal{L} \\ \nu \end{bmatrix} \tag{57}$$

The updated values of x, p, q, β, α, γ, and λ are computed using the primal and dual step method proposed by [52]. This method is utilized in the new iterations and thus, determines the values of the aforementioned variables as

$$\begin{cases} [x^{(k)}, p^{(k)}, q^{(k)}]^T + \Delta_P[\mathrm{d}x, \mathrm{d}p, \mathrm{d}q]^T \\ [\alpha^{(k)}, \beta^{(k)}, \lambda^{(k)}, \gamma^{(k)}]^T + \Delta_D[\mathrm{d}\alpha, \mathrm{d}\beta, \mathrm{d}\lambda, \mathrm{d}\gamma]^T \end{cases} \tag{58}$$

where

$$\begin{cases} \Delta_P = 0.9995 \min \\ \left(\min\left(-\dfrac{q_i}{dq_i} : q_i < 0; -\dfrac{p_i}{dp_i} : p_i < 0\right), 1\right) \\ \Delta_D = 0.9995 \min \\ \left(\min\left(-\dfrac{\gamma_i}{d\gamma_i} : \gamma_i < 0; -\dfrac{\lambda_i}{d\lambda_i} : \lambda_{2i} < 0\right), 1\right) \end{cases}$$

B. ADAPTIVE GCC

To automatically utilize the nature of the data and the defined GCC criterion, we introduce a new optimization layer in this section. This layer computes the optimal value of c for each static snapshot, taking into account the GGD curve which can

range from a Laplace to a uniform distribution, as shown in Fig. 1. The GCC will adapt to the most suitable distribution at any given time.

Consider $F(x)$, as defined in equation (24). We compute the derivative of $F(x)$ with respect to c, while fixing ς, to obtain:

$$\frac{\partial F(x)}{\partial c} = \frac{F(x)}{m}\left(A\,\frac{\partial B}{\partial c} + \frac{\partial A}{\partial c}\,B\right) \tag{59}$$

where

$$A = \frac{c}{2\,\omega\,\Gamma\left(\frac{1}{c}\right)}, \quad B(x) = \exp\left(-\left|\frac{r(x)}{\omega}\right|^{c}\right)$$

$$\frac{\partial A}{\partial c} = \frac{\frac{1}{2}\,\omega\,\Gamma\left(\frac{1}{c}\right) - c\,\frac{\partial\left(\omega\,\Gamma\left(\frac{1}{c}\right)\right)}{\partial c}}{\left(\omega\,\Gamma(\frac{1}{c})\right)^{2}}$$

$$\frac{\partial\left(\omega\,\Gamma\left(\frac{1}{c}\right)\right)}{\partial c} = \frac{\partial \omega}{\partial c}\,\Gamma\left(\frac{1}{c}\right) + \omega\,\psi\left(\frac{1}{c}\right)$$

$$\frac{\partial \omega}{\partial c} = \frac{\varsigma\left(\frac{3\,\Gamma\left(\frac{1}{c}\right)\psi\left(\frac{3}{c}\right)}{c^{2}\,\Gamma\left(\frac{3}{c}\right)} - \frac{\Gamma\left(\frac{1}{c}\right)\psi\left(\frac{1}{c}\right)}{c^{2}\,\Gamma\left(\frac{3}{c}\right)}\right)}{\left(\frac{\Gamma\left(\frac{1}{c}\right)}{\Gamma\left(\frac{3}{c}\right)}\right)^{\frac{1}{2}}}$$

$$\frac{\partial B}{\partial c} = -B\Bigg(\log(\sigma_1)\sigma_1^{c} + \frac{c\,|r(x)|\,\sigma_1^{c-1}\Gamma\left(\frac{1}{c}\right)\psi\left(\frac{1}{c}\right) sign\left(\Gamma\left(\frac{1}{c}\right)\right)}{2\,c^{2}\,\sigma_2^{\frac{3}{2}}|\varsigma|\left|\Gamma\left(\frac{3}{c}\right)\right|}\Bigg) + B\left(\frac{3\,c\,\Gamma\left(\frac{3}{c}\right)|r(x)|\,\left|\Gamma\left(\frac{1}{c}\right)\right|\psi\left(\frac{3}{c}\right) sign\left(\Gamma\left(\frac{3}{c}\right)\right)}{2\,c^{2}\,\sigma_2^{\frac{3}{2}}|\varsigma|\left|\Gamma\left(\frac{3}{c}\right)\right|^{2}}\right)$$

$$\sigma 1 = \frac{|r(x)|}{|\varsigma|\sqrt{\sigma_2}}, \quad \sigma_2 = \left|\frac{\Gamma\left(\frac{1}{c}\right)}{\Gamma\left(\frac{3}{c}\right)}\right|, \quad \psi\left(\frac{1}{c}\right) = \frac{\partial\Gamma\left(\frac{1}{c}\right)}{\partial c}$$

Therefore, the residual $r(x)$, which appears in the derivative of the objective function with respect to c in (59), is substituted by the new residual v, computed in (56). For multiple values of c, c_{new} is determined when ${}^{\partial F}/_{\partial c} = 0$, and checked to confirm if it matches the old value of c. If they do not match, the AMCC algorithm updates c for a new iteration.

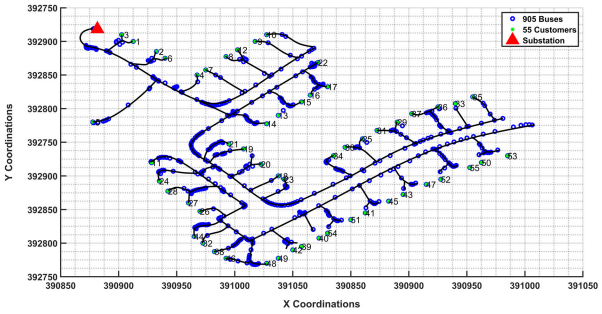


FIGURE 3. Map diagram of the European LV feeder.

The overall algorithm for the proposed method is presented in Table 2.

IV. CASE STUDY

In this section, we present a case study that highlights a real low voltage network and its associated data. Using Matlab software, we carry out state estimation (SE) simulations employing different algorithms and assess the accuracy and time efficiency of each method. We then provide a comprehensive discussion of the simulation outcomes.

A. NETWORK DESCRIPTION

Real data from a network located in Kent, England, is utilized to validate the SE methods presented in this study [35]. In addition to the GE challenge, this network is a radial distribution grid, which poses a greater challenge for the SE algorithms. The low voltage radial distribution map diagram of the LV feeder is shown in Fig. 3, and it consists of 55 household customers, a distribution substation, 905 buses, a frequency of 50Hz, and a 3-phase $\triangle/Y$-grounded transformer (11kV to 0.416kV phase to phase). During simulation, 11kV is equal to 1.05p.u. The customer load profile data is collected at a 1-minute resolution over one day and is obtained from [54]. For simplicity, the number of nodes is reduced to 22 buses by lumping many customer loads in the edited network presented in Fig. 4.

B. LOAD PROFILE DATA

In this scenario, Bus 1 serves as the feeder bus (slack bus type), while buses 5 through 13 are considered zero-injection buses. All other buses are categorized as load buses (PQ buses type). In this setup, the real power (MW) and reactive power (MVAr) demands were measured over a 24-hour period.

An example of the load profile data for bus 2 is presented in Fig. 5, with data points recorded at one-minute intervals.

To illustrate the behavior of proposed SE algorithm, the following general methodology is followed:

1) Collect load profile data over a 24-hour period and use the Matpower tool [53] to solve the steady-state power flow problem. This will generate a set of measurement vectors comprised of various parameters, including the voltage amplitude of the first bus, as well as the active

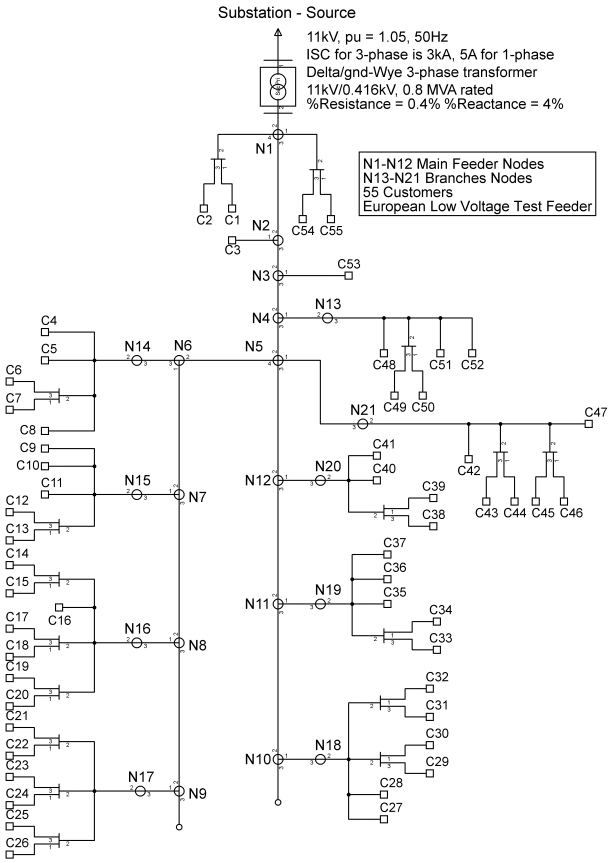


FIGURE 4. One-line diagram of the LV network [41].

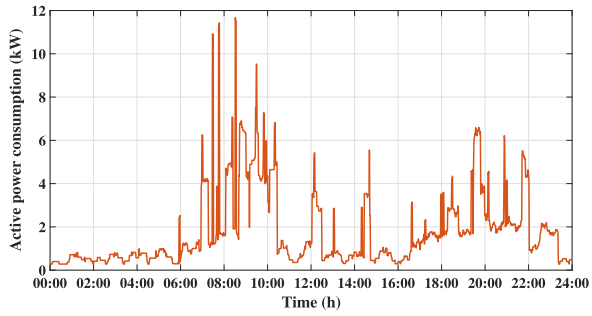


FIGURE 5. Load profile data consumption in the bus 2.

and reactive power demands of 22 buses (44 measurements), and the active and reactive power flow between buses 1-2 and 2-3 (4 measurements). In this study, a measurement vector comprising 49 elements is used, which are known as true measurements (z_{true}) and were recorded at a resolution of 1 minute.

2) Add 10% Gaussian noise (N) to the true measurements, except for the feeder bus voltage, to generate contaminated measurement signals denoted as $z_{noisy} = z_{true} + N$ [56].
3) To account for the presence of GE in z_{noisy}, a randomly (see next subsection) value in the order of -1 is added. Monte Carlo simulations are employed to simulate 96 snapshots, each spanning 15 minutes over a 24-hour period. This process is repeated 10 times, resulting in a total of 960 iterations, in order to obtain accurate time processing and RMSE mean values for each method.

TABLE 3. Branches data.

From bus	To bus	R (pu)	X (pu)
1	2	0.0016	0.0012
2	3	0.0090	0.0068
3	4	0.0025	0.0019
4	5	0.0026	0.0020
5	6	0.0010	0.0007
5	14	0.0063	0.0047
6	22	0.0095	0.0072
6	7	0.0092	0.0070
6	13	0.0053	0.0040
7	15	0.0024	0.0018
7	8	0.0060	0.0046
8	16	0.0043	0.0033
8	9	0.0058	0.0044
9	17	0.0048	0.0036
9	10	0.0057	0.0043
10	18	0.0039	0.0029
13	21	0.0058	0.0044
13	12	0.0032	0.0024
12	20	0.0038	0.0029
12	11	0.0157	0.0119
11	19	0.0071	0.0054

Table 3 presents the branch data for the low-voltage network used in the "Matpower" tool. The values of impedance (R) and inductance (X) were calculated using the aggregation method.

C. STUDY DATA CASES

The mean RMSE of the state variables (voltage angle θ and amplitude V) were computed for 22 buses over a period of 24 hours and compared for each SE method under various data scenarios. The performance of the SE methods was evaluated based on the one-day RMSE, as shown below:

$$\text{RMSE} = \sqrt{\frac{1}{u}\sum_{l=1}^{u}(\hat{x}_l - x_l)^2} \tag{60}$$

where $\hat{x}$ is the estimated value, x is the true value, and u is the number of static snapshots per day [9].

A comparison was also made between the WLS and the studied SE methods. The same weighted covariance R_{ii} was used in all the WLS, Cauchy, WLAV, and Hachtel methods.

This section presents the measurement data used in the study. Aleatory GE was injected into the noisy measurements (z_{noisy}), and the percentage of the GE case study was calculated from the number of GEs in the total measurements data vector length (49). Three measurement scenarios were studied and tested:

1) SCENARIO OF 4% GE

Two GEs are added aleatory to z_{noisy}. Fig. 6 displays the voltage amplitude and angle estimates for bus 6 obtained using the proposed SE methods over a 24-hour period and

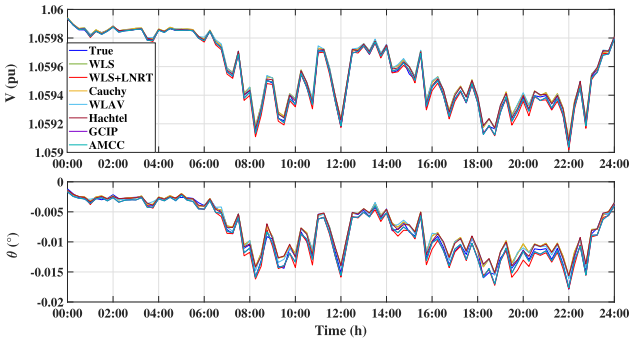


FIGURE 6. Estimation of the voltage amplitude and angle during 24*h* in the bus 6, for the case of 4% GE.

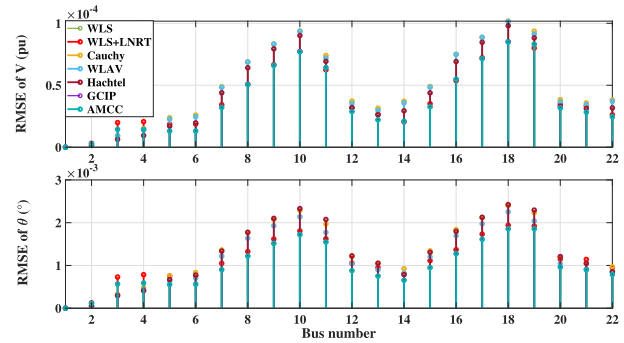


FIGURE 7. Voltage amplitude and angle RMSE in each bus, for the case of 4% GE.

compares them with the true measurements. Fig. 7 displays the obtained RMSE of each SE methods for each bus.

Table 4 displays the RMSE values for all buses and static snapshots during a 24*h* period. The accuracy of the estimates for each algorithm is listed in ascending order. In the case of 4% GE, the AMCC method exhibits better accuracy compared to other studied SE methods. Table 5 shows the mean RMSE difference between the WLS method and the studied SE methods. In the case of WLS-Cauchy, the mean RMSE difference for V is in order of -10 (digit) for V and -9 for θ. For WLAV and Hachtel methods, the mean RMSE difference is in order of -6 and -5, respectively, for V, and -4 for θ.

The shape parameter of the kernel function in the GCIP method is chosen as $c = 13.4$ for optimal performance, as reported in [34]. However, the AMCC approach optimizes the suitable value of c for each set of measurement data, making it an effective algorithm for static snapshots. In reality, the optimal value of c varies for each data scenario to achieve the highest accuracy.

In this case study, the AMCC, GCIP, and LNRT methods were found to be the most accurate SE methods, with mean RMSE differences of approximately -4 for V (excluding LNRT, which had a difference of -5) and -3 for θ. Notably, the accuracy difference between the AMCC and LNRT methods was significant at around -5. Furthermore, the AMCC method outperformed even the GCIP model in terms of accuracy, likely due to its adaptive c parameter, which is approximately -8.

TABLE 4. Mean RMSE of the residuals for all buses of the SE methods in 4% GE case.

Method	RMSE of V (pu)	RMSE of θ (°)
WLS	0.4862794×10^{-4}	0.1283484×10^{-2}
Cauchy	0.4862793×10^{-4}	0.1283483×10^{-2}
WLAV	0.4777431×10^{-4}	0.1158973×10^{-2}
Hachtel	0.4402787×10^{-4}	0.1270748×10^{-2}
WLS+LNRT	0.3960839×10^{-4}	0.1122920×10^{-2}
GCIP	0.3778188×10^{-4}	0.0989366×10^{-2}
AMCC	0.3775196×10^{-4}	0.0989182×10^{-2}

TABLE 5. Difference of the Mean RMSE between studied SE methods and the WLS in 4% GE case.

Method	RMSE of V (pu)	RMSE of θ (°)
WLS-Cauchy	0.1531856×10^{-10}	0.7240420×10^{-9}
WLS-WLAV	0.8536380×10^{-6}	0.1245103×10^{-4}
WLS-Hachtel	0.4600076×10^{-5}	0.1273554×10^{-4}
WLS-LNRT	0.9019553×10^{-5}	0.1605633×10^{-3}
WLS-GCIP	0.1084606×10^{-4}	0.2941181×10^{-3}
WLS-AMCC	0.1085598×10^{-4}	0.2943014×10^{-3}

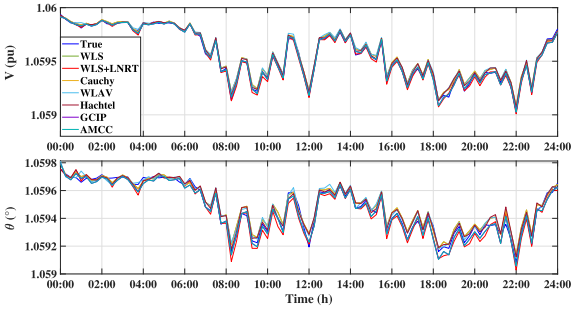


FIGURE 8. Estimation of the voltage amplitude and angle during 24*h* in bus 6, for the case of 8% GE.

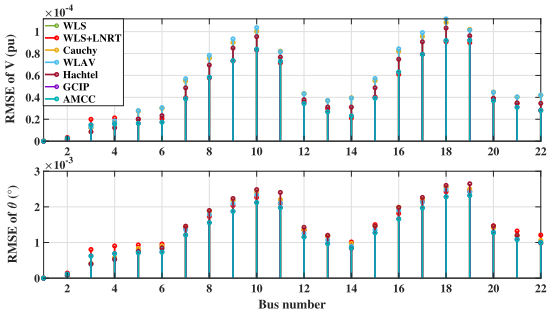


FIGURE 9. Voltage amplitude and angle RMSE in each bus, for the case of 8% GE.

2) SCENARIO OF 8% GE

The data was contaminated with 4 GE, and the estimation of state variables at bus 6 using the SE methods studied were displayed in Fig. 8 for the duration of 24 hours. The voltage amplitude and angle RMSE of each method for every bus during this period are also shown in Fig. 9.

The mean RMSE residuals for all buses are presented in Table 6, showing good accuracy for the AMCC in this case study as well. The studied methods are classified in order of

TABLE 6. Mean RMSE of the residuals for all buses of the SE methods in 8% GE case.

Method	RMSE of V (pu)	RMSE of θ $(^\circ)$
WLS	0.5372829×10^{-4}	0.1399925×10^{-2}
Cauchy	0.5372828×10^{-4}	0.1399925×10^{-2}
WLAV	0.5480072×10^{-4}	0.1323070×10^{-2}
Hachtel	0.4830854×10^{-4}	0.1316924×10^{-2}
WLS+LNRT	0.4359939×10^{-4}	0.1316737×10^{-2}
GCIP	0.4272107×10^{-4}	0.1247206×10^{-2}
AMCC	0.4271045×10^{-4}	0.1246534×10^{-2}

TABLE 7. Difference of the mean RMSE between studied SE methods and the WLS in 8% GE case.

Method	RMSE of V (pu)	RMSE of θ $(^\circ)$
WLS-Cauchy	0.7836902×10^{-11}	0.6759557×10^{-9}
WLS-WLAV	0.1072433×10^{-5}	0.7685513×10^{-4}
WLS-Hachtel	0.5419754×10^{-5}	0.1699908×10^{-4}
WLS-LNRT	0.1012890×10^{-4}	0.1681129×10^{-4}
WLS-GCIP	0.1100721×10^{-4}	0.1527192×10^{-3}
WLS-AMCC	0.1101783×10^{-4}	0.1533916×10^{-3}

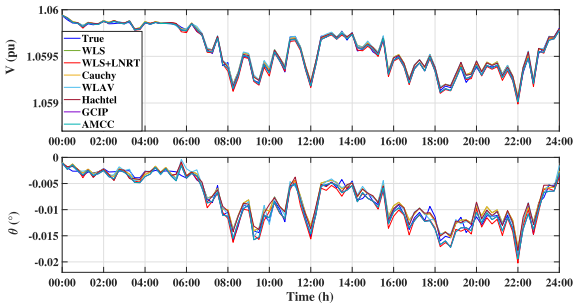


FIGURE 10. Estimation of the voltage amplitude and angle during 24*h* in the bus 6, for the case of 16% GE.

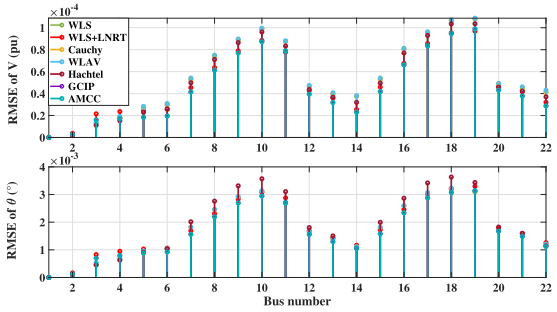


FIGURE 11. Voltage amplitude and angle RMSE in each bus, for the case of 16% GE.

accuracy, as shown in Table 7. The same comments made in the previous data case study are applicable, where the proposed AMCC method outperforms other SE models in terms of accuracy, at the same levels. Some slight differences have occurred in the mean RMSE difference, such as with the WLS-LNRT methods, which decreased to an order of -4 for V and increased to an order of -4 for θ.

3) SCENARIO OF 16% GE

An extreme case is examined, where 8 GEs are included in the noisy measurements. The estimation of state variables at bus 6 for each of the studied methods is presented in comparison with the true measurements in Fig. 10. Furthermore, the computed RMSE for each bus over a period of 24 hours is shown in Fig. 11.

TABLE 8. Mean RMSE of the residuals for all buses of the SE methods in 16% GE case.

Method	RMSE of V (pu)	RMSE of θ $(^\circ)$
WLS	0.5399577×10^{-4}	0.1794899×10^{-2}
Cauchy	0.5399576×10^{-4}	0.1794898×10^{-2}
WLAV	0.5328211×10^{-4}	0.1681156×10^{-2}
Hachtel	0.5013030×10^{-4}	0.1880456×10^{-2}
WLS+LNRT	0.4363538×10^{-4}	0.1726677×10^{-2}
GCIP	0.4205326×10^{-4}	0.1617211×10^{-2}
AMCC	0.4203076×10^{-4}	0.1616537×10^{-2}

TABLE 9. Difference of the mean RMSE between studied SE methods and the WLS in 16% GE case.

Method	RMSE of V (pu)	RMSE of θ $(^\circ)$
WLS-Cauchy	0.1185017×10^{-10}	0.6196838×10^{-9}
WLS-WLAV	0.7136529×10^{-6}	0.1137431×10^{-3}
WLS-Hachtel	0.3865471×10^{-5}	0.8555761×10^{-4}
WLS-LNRT	0.1036038×10^{-4}	0.6822131×10^{-4}
WLS-GCIP	0.1194250×10^{-4}	0.1776876×10^{-3}
WLS-AMCC	0.1196500×10^{-4}	0.1783613×10^{-3}

Table 8 presents the mean RMSE residuals for all buses over a period of 24 hours, arranged in decreasing order of mean RMSE. Similarly, Table 9 shows the difference in mean RMSE between the WLS method and other SE methods. While slight differences are noticed, the proposed AMCC method has demonstrated its superiority with a difference in order of -4 compared to the WLS, -5 compared to the LNRT, and -7 compared to the GCIP.

4) TIME PERFORMANCE

The simulation was run on a machine equipped with an Intel(R) Core(TM) i7-8750H processor, with a clock speed of 2.20-2.21GHz, and 16GB of RAM installed. No other tasks were running in parallel during the simulation. Table 10 displays the average time required for each algorithm to complete one iteration per static snapshot for the three measurement data scenarios. The algorithms are ranked by speed, with the Hachtel method taking the longest time to process, while the GCIP is the fastest.

Our results indicate that all three methods, namely AMCC, GCIP, and LNRT, exhibit good accuracy. However, it is worth noting that the AMCC method has a longer processing time compared to the GCIP method, which is understandable since it involves repeating the GCIP process to find the optimal c for each data set scenario.

Furthermore, the AMCC method requires more time for iteration compared to the LNRT method. This is reasonable as the LNRT method simplifies calculations by deleting a certain number of suspected measurements, allowing it to achieve the required tolerance in a shorter processing time. However, as we mentioned earlier, the LNRT method presents a challenge in terms of network observability, which demands a well-tuned threshold depending on the network's sensitivity.

TABLE 10. Mean iteration time needed during all static snapshots for each algorithm method in *ms*.

SE methods	4% GE	8% GE	16% GE
Hachtel	28.336	24.573	27.814
WLAV	11.754	11.622	11.821
AMCC	9.249	9.635	9.671
WLS	5.375	5.299	5.718
Cauchy	5.443	5.216	5.517
LNRT	4.001	3.978	4.536
GCIP	3.284	3.332	3.340

Additionally, the LNRT method sets a limit on the suspected measurements to be deleted, and hence, the remaining GEs are used as data set elements. Without proper tuning, the LNRT method may not provide good accuracy as per our case study.

The WLAV method necessitates more processing time because its state estimation formulation is based on the IPM. This entails computing more parameters in the form of slack variables. On the other hand, the Hachtel method involves computing other parameters such as Lagrange multipliers, which also adds to its processing time requirements.

V. CONCLUSION

The presence of multiple GE can cause inaccurate estimation of system states and mislead system operators. To tackle this challenge, several robust SE methods have been proposed and presented.

In this study, we discuss the importance of robust SE and error elimination. Additionally, we present a comparison between several algorithms, including least squares associated with the LNRT, Cauchy, least absolute value, and Hachtel SE methods.

Another SE method, named GCIP, is also presented in this work. GCIP uses a cost function based on generalized correntropy, with a high shape parameter c, and an interior point solving approach. However, to address the drawbacks of GCIP, we propose an accurate and robust SE algorithm for static snapshots, called AMCC. In this algorithm, c is iteratively estimated for each data scenario.

For real case data scenario, the robustness of the SE methods towards multiple GE in aleatory points of the network is addressed. Specifically, three data scenarios were adopted in the test-bed simulation, where the measurements data set was corrupted with 4%, 8%, and 16% of GE, respectively. Monte Carlo simulations are utilized to accurately compute the mean values of RMSE and processing time for each method.

The SE methods studied in this research were validated using real low-voltage network data from Kent, UK. The results indicate that the proposed AMCC method is robust against GE in the addressed data scenarios and performs better than the WLS, Cauchy, WLAV, Hachtel, WLS+LNRT, and GCIP methods.

The proposed method has demonstrated its superiority in terms of accuracy and flexibility for successive static snapshots. However, we must also acknowledge its drawback concerning time processing (for our case in the order of 9ms) and address it in future work.

REFERENCES

[1] A. Abur and A. G. Expósito, *Power System State Estimation*. Boca Raton, FL, USA: CRC Press, 2004, doi: 10.1201/9780203913673.

[2] N. G. Bretas and A. S. Bretas, "A two steps procedure in state estimation gross error detection, identification, and correction," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 484–490, Dec. 2015, doi: 10.1016/j.ijepes.2015.05.044.

[3] 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, doi: 10.1109/TSG.2017.2761452.

[4] M. M. Saha, J. Izykowski, and E. Rosolowski, *Fault Location on Power Networks* (Power Systems). London, U.K.: Springer, 2010, doi: 10.1007/978-1-84882-886-5.

[5] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," *Int. J. Elect. Power Energy Syst.*, vol. 63, pp. 473–484, Dec. 2014, doi: 10.1016/j.ijepes.2014.06.025.

[6] A. Abdel-Majeed and M. Braun, "Low voltage system state estimation using smart meters," in *Proc. 47th Int. Universities Power Eng. Conf. (UPEC)*, Sep. 2012, pp. 1–6, doi: 10.1109/UPEC.2012.6398598.

[7] J. Liu, J. Tang, F. Ponci, A. Monti, C. Muscas, and P. A. Pegoraro, "Trade-offs in PMU deployment for state estimation in active distribution grids," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 915–924, Jun. 2012, doi: 10.1109/TSG.2012.2191578.

[8] I. Dzafic, M. Gilles, R. A. Jabr, B. C. Pal, and S. Henselmeyer, "Real time estimation of loads in radial and unsymmetrical three-phase distribution networks," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4839–4848, Nov. 2013, doi: 10.1109/TPWRS.2013.2277888.

[9] F. Schweppe and J. Wildes, "Power system static-state estimation, part I: Exact model," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 120–125, Jan. 1970, doi: 10.1109/TPAS.1970.292678.

[10] S. Basumallik, R. Ma, and S. Eftekharnejad, "Packet-data anomaly detection in PMU-based state estimator using convolutional neural network," *Int. J. Electr. Power Energy Syst.*, vol. 107, pp. 690–702, May 2019, doi: 10.1016/j.ijepes.2018.11.013.

[11] R. Singh, B. C. Pal, and R. A. Jabr, "Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement," *IET Gener., Transmiss. Distrib.*, vol. 4, no. 1, p. 50, 2010, doi: 10.1049/iet-gtd.2009.0167.

[12] A. Mouco and A. Abur, "A robust state estimator for power systems with HVDC components," in *Proc. North Amer. Power Symp. (NAPS)*, Morgantown, WV, USA, Sep. 2017, pp. 1–5, doi: 10.1109/NAPS.2017.8107197.

[13] 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, doi: 10.1109/ACCESS.2018.2853664.

[14] O. Alsac, N. Vempati, B. Stott, and A. Monticelli, "Generalized state estimation," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1069–1075, Aug. 1998, doi: 10.1109/59.709101.

[15] P. J. Huber, *Robust Statistics: Huber/Robust Statistics*. Hoboken, NJ, USA: Wiley, 1981, doi: 10.1002/0471725250.

[16] M. E. Baran and A. W. Kelley, "A branch-current-based state estimation method for distribution systems," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 483–491, Feb. 1995, doi: 10.1109/59.373974.

[17] Y. Chen, H. Chen, Y. Jiao, J. Ma, and Y. Lin, "Data-driven robust state estimation through off-line learning and on-line matching," *J. Modern Power Syst. Clean Energy*, vol. 9, no. 4, pp. 897–909, 2021, doi: 10.35833/MPCE.2020.000835.

[18] Y. Chen, C. Wu, and J. Qi, "Data-driven power flow method based on exact linear regression equations," *J. Modern Power Syst. Clean Energy*, vol. 10, no. 3, pp. 800–804, 2022, doi: 10.35833/MPCE.2020.000738.

[19] D. M. Falcao and S. M. de Assis, "Linear programming state estimation: Error analysis and gross error identification," *IEEE Trans. Power Syst.*, vol. PS-3, no. 3, pp. 809–815, Aug. 1988, doi: 10.1109/59.14526.

[20] N. G. Bretas, A. S. Bretas, and S. A. Piereti, "Innovation concept for measurement gross error detection and identification in power system state estimation," *IET Gener., Transmiss. Distrib.*, vol. 5, no. 6, p. 603, 2011, doi: 10.1049/iet-gtd.2010.0459.

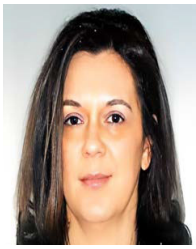
[21] I. D. Melo and M. P. Antunes, "Bad data correction in harmonic state estimation for power distribution systems: An approach based on generalised pattern search algorithm," *Electr. Power Syst. Res.*, vol. 204, Mar. 2022, Art. no. 107684, doi: 10.1016/j.epsr.2021.107684.
[22] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Two deterministic half-quadratic regularization algorithms for computed imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 2, Austin, TX, USA, Sep. 1994, pp. 168–172, doi: 10.1109/ICIP.1994.413553.
[23] L. Catania and A. Luati, "Robust estimation of a location parameter with the integrated Hogg function," *Statist. Probab. Lett.*, vol. 164, Sep. 2020, Art. no. 108812, doi: 10.1016/j.spl.2020.108812.
[24] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image Vis. Comput.*, vol. 15, no. 1, pp. 59–76, Jan. 1997, doi: 10.1016/S0262-8856(96)01112-2.
[25] B. C. Penney, M. A. King, R. B. Schwinger, S. P. Baker, and P. W. Doherty, "Modifying constrained least-squares restoration for application to single photon emission computed tomography projection images: Application of least-squares restoration to SPECT images," *Med. Phys.*, vol. 15, no. 3, pp. 334–342, May 1988, doi: 10.1118/1.596227.
[26] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comput. Vis. Image Understand.*, vol. 63, no. 1, pp. 75–104, Jan. 1996, doi: 10.1006/cviu.1996.0006.
[27] J. E. Dennis and R. E. Welsch, "Techniques for nonlinear least squares and robust regression," *Commun. Statist.-Simul. Comput.*, vol. 7, no. 4, pp. 345–359, 1978, doi: 10.1080/03610917808812083.
[28] Y. Chen, J. Ma, P. Zhang, F. Liu, and S. Mei, "Robust state estimator based on maximum exponential absolute value," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1537–1544, Jul. 2017, doi: 10.1109/TSG.2015.2485280.
[29] B. Chen, Y. Xie, Z. Li, Y. Li, and P. Ren, "Asymmetric correntropy for robust adaptive filtering," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 3, pp. 1922–1926, Mar. 2022, doi: 10.1109/TCSII.2021.3122283.
[30] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007, doi: 10.1109/TSP.2007.896065.
[31] Y. Chen, F. Liu, S. Mei, and J. Ma, "Toward adaptive robust state estimation based on MCC by using the generalized Gaussian density as kernel functions," *Int. J. Electr. Power Energy Syst.*, vol. 71, pp. 297–304, Oct. 2015, doi: 10.1016/j.ijepes.2015.03.011.
[32] S. Pesteh, H. Moayyed, V. Miranda, J. Pereira, V. Freitas, A. S. Costa, and J. B. A. London, "A new interior point solver with generalized correntropy for multiple gross error suppression in state estimation," *Electr. Power Syst. Res.*, vol. 176, Nov. 2019, Art. no. 105937, doi: 10.1016/j.epsr.2019.105937.
[33] H. Moayyed, D. Ghaderyan, Y. Boukili, and A. P. Aguiar, "Accelerated generalized correntropy interior point method in power system state estimation," in *Proc. CONTROLO*. Cham, Switzerland: Springer, 2021, pp. 658–667, doi: 10.1007/978-3-030-58653-9_63.
[34] S. Pesteh, H. Moayyed, and V. Miranda, "Favorable properties of interior point method and generalized correntropy in power system state estimation," *Electr. Power Syst. Res.*, vol. 178, Jan. 2020, Art. no. 106035, doi: 10.1016/j.epsr.2019.106035.
[35] A. N. Espinosa, "Low voltage networks models and low carbon technology profiles," Dept. Elect. Electron. Eng., Univ. Manchester, Manchester, U.K., Tech. Rep., Jun. 2015.
[36] A. A. Prasad, R. A. Taylor, and M. Kay, "Assessment of solar and wind resource synergy in Australia," *Appl. Energy*, vol. 190, pp. 354–367, Mar. 2017, doi: 10.1016/j.apenergy.2016.12.135.
[37] J. Aldrich, "RA Fisher and the making of maximum likelihood 1912–1922," *Stat. Sci.*, vol. 12, no. 3, pp. 162–176, Sep. 1997, doi: 10.1214/ss/1030037906.
[38] M. G. Kendall, "Studies in the history of probability and statistics: XI. Daniel Bernoulli on maximum likelihood," *Biometrika*, vol. 48, nos. 1–2, pp. 1–2, 1961, doi: 10.1093/biomet/48.1-2.1.
[39] A. Hald, "On the history of maximum likelihood in relation to inverse probability and least squares," *Statist. Sci.*, vol. 14, no. 2, May 1999, doi: 10.1214/ss/1009212248.
[40] Y. Chen, "Weighted-Least-Square (WLS) state estimation," Chem. Rubber Company (CRC) Press, Boca Raton, FL, USA, Tech. Rep., Dec. 2015.
[41] 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*, Milan, Italy, Jun. 2019, pp. 1–6, doi: 10.1109/PTC.2019.8810929.
[42] G. N. Korres, "A robust method for equality constrained state estimation," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 305–314, May 2002, doi: 10.1109/TPWRS.2002.1007897.
[43] M. Göl 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, doi: 10.1109/TSG.2014.2302213.
[44] S. M. Stigler, "Studies in the history of probability and statistics. XXXIII Cauchy and the witch of Agnesi: An historical note on the Cauchy distribution," *Biometrika*, vol. 61, no. 2, pp. 375–380, 1974, doi: 10.1093/biomet/61.2.375.
[45] W. J. J. Rey, *Introduction to Robust and Quasi-Robust Statistical Methods*. Berlin, Germany: Springer, 1983, doi: 10.1007/978-3-642-69389-2.
[46] A. Gjelsvik, S. Aam, and L. Holten, "Hachtel's augmented matrix method—A rapid method improving numerical stability in power system static state estimation," *IEEE Trans. Power Apparatus Syst.*, vol. PAS-104, no. 11, pp. 2987–2993, Nov. 1985, doi: 10.1109/TPAS.1985.318939.
[47] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, "Inexact Newton methods," *SIAM J. Numer. Anal.*, vol. 19, no. 2, pp. 400–408, 1982, doi: 10.1137/0719025.
[48] A. L. Langner and A. Abur, "Formulation of three-phase state estimation problem using a virtual reference," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 214–223, Jan. 2021, doi: 10.1109/TPWRS.2020.3004076.
[49] F. Balogh and F. RÈnyi-Vámos, "Some problems of pyelonephritis," *Urologia J.*, vol. 43, no. 6, pp. 663–667, Dec. 1976, doi: 10.1177/039156037604300622.
[50] W. Wu, Y. Guo, B. Zhang, A. Bose, and S. Hongbin, "Robust state estimation method based on maximum exponential square," *IET Gener., Transmiss. Distrib.*, vol. 5, no. 11, p. 1165, 2011, doi: 10.1049/iet-gtd.2011.0100.
[51] A. S. El-Bakry, R. A. Tapia, T. Tsuchiya, and Y. Zhang, "On the formulation and theory of the Newton interior-point method for nonlinear programming," *J. Optim. Theory Appl.*, vol. 89, no. 3, pp. 507–541, Jun. 1996, doi: 10.1007/BF02275347.
[52] H. Wei, H. Sasaki, J. Kubokawa, and R. Yokoyama, "An interior point method for power system weighted nonlinear L/sub 1/norm static state estimation," *IEEE Trans. Power Syst.*, vol. 13, no. 2, pp. 617–623, May 1998, doi: 10.1109/59.667390.
[53] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011, doi: 10.1109/TPWRS.2010.2051168.
[54] IEEE PES Distribution System Analysis Subcommittee Test Feeders Working Group. *IEEE Distribution Test Feeders*. Accessed: Jan. 11, 2020. [Online]. Available: https://cmte.ieee.org/pes-testfeeders/
[55] P. Grahn, V. Briggner, L. Johansson, D. Babazadeh, L. Nordstrom, and D. Babazadeh, "Centralized versus distributed state estimation for hybrid AC/HVDC grid," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Eur. (ISGT-Europe)*, Sep. 2017, pp. 1–6, doi: 10.1109/ISGTEurope.2017.8260167.
[56] 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, Feb. 2010.

YASSINE BOUKILI received the B.S. degree in industrial engineering and the M.S. degree in mechatronics from the University of Sidi Mohamed Ben Abdellah (USMBA), Fez, Morocco, in 2011 and 2014, respectively. Furthermore, he began his Ph.D. studies at the Faculty of Engineering, University of Porto (FEUP), Portugal in September 2018. Currently, he is a Researcher with SYSTEC-FEUP and has previously worked on the intelligent management of renewable energy systems project, funded by the National Center for Scientific and Technical Research, Morocco. His research interests include state estimation, optimization and smart grids, nonlinear and optimal control, and the integration of renewable energy resources into conventional power systems.

MOTAZ AYIAD received the B.S. degree in mechatronics engineering from Al-Azhar University-Gaza, in 2015, the M.S. degree in future power networks from Imperial College London, in 2017, and the Ph.D. degree in electrical engineering from the University of Porto, in 2022. He is currently with the Research and Development Department, P&C, Efacec, Portugal. He was a member of an EU-funded project under Marie Curie Actions called InnoDC. His research interests include power system state estimation, mainly hybrid HV-ac/dc networks. In addition, his research covers the deployment of SCADA systems and P&C IEDs for ac/dc grids.

HAMED MOAYYED received the M.Sc. degree in mathematics from the K. N. Toosi University of Technology, Tehran, Iran, and the Ph.D. degree in optoelectronics from the University of Porto, Portugal, in 2016. He is currently researching the applications of deep neural networks in modern power systems. He is also experienced in mathematical modeling, computer simulations, optimization, numerical analysis, and machine learning techniques. His main research interests include smart energy systems, power systems, state estimation, and power system cyber security.

A. PEDRO AGUIAR (Senior Member, IEEE) received the bachelor's, M.S., and Ph.D. degrees in electrical and computer engineering (ECE) from Instituto Superior Técnico (IST), University of Lisbon, Portugal, in 1994, 1998, and 2002, respectively. He is currently a Professor with the Department of Electrical and Computer Engineering, Faculty of Engineering, University of Porto (FEUP). His research interests include control systems, robotics, and signal processing, with a particular focus on mobile networked cyber-physical systems; motion planning, guidance, navigation, and control of single and multiple cooperative/coordinated autonomous robotic vehicles; nonlinear control and estimation theory; optimization-based and optimal control; networked control; the integration of machine learning with feedback control; and large-scale distributed systems.



ZITA VALE (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering and the Habilitation degree from the University of Porto, Porto, Portugal, in 1993. She is currently a Full Professor with the Polytechnic Institute of Porto, Porto. She is also the Director of the GECAD-Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development. Her research interests include artificial intelligence applications, smart grids, electricity market, demand response, electric vehicles, and renewable energy sources.