WLS-based State Estimation for Unobservable Distribution Grids through Allocation Factors Evaluation

Marco Pau, *Member, IEEE*, Paolo Attilio Pegoraro, *Senior Member, IEEE*

***Abstract*—Real-time monitoring of distribution grids is essential to handle the complex operation of modern electric systems. One of the main challenges for the deployment of reliable monitoring solutions at distribution level is the scarcity of measurement instrumentation in the field. Despite the proposal of some approaches to deal with the under-determined system caused by the low number of meters, existing solutions are not yet able to guarantee a level of simplicity and trustworthiness similar to the one of conventional Weighted Least Squares (WLS) estimators adopted in transmission systems. This paper aims at filling this gap by presenting a WLS-based estimator able to work with only very few meters, in scenarios typically considered as unobservable, and without the need to employ pseudo-measurements. The proposed method relies on the use of allocation factors and requires only minor modifications with respect to the conventional WLS, thus offering the benefits associated to the use of a well-known and mature state estimation formulation. Simulations performed on an unbalanced IEEE test-grid highlight the performance and advantages of the proposed estimator, proving its suitability for the monitoring of poorly instrumented distribution grids.**

***Index Terms*—Power distribution networks, Power system measurements, Observability, State Estimation, System Awareness, Weighted Least Squares, Voltage measurement, Power measurement.**

I. Introduction

Distribution System Operators (DSOs) are currently digitalizing their Medium and Low Voltage (MV and LV) grids and equipping their control rooms with the software functionalities needed for the smart management and automation of the network [1]. Among these functionalities, monitoring the distribution system (DS) via ad hoc State Estimation (SE) methods is an impelling requirement [2]. SE unlocks awareness about the real-time operating conditions, which is crucial to enable contingency analysis and to ensure optimal control of flexible resources, such as new types of loads (e.g., electric vehicles) and Distributed Generation (DG).

The design of reliable SE algorithms for DS monitoring is not a trivial task. Several challenges hinder the easy deployment of SE at distribution level and prevent the adoption of SE solutions commonly used in transmission [3]. The most critical issue is the scarcity of measurement devices in the field. Also due to their very large size (in topological terms, i.e., number of nodes and branches), distribution grids are generally non-observable, namely the SE model results into an under-determined system where only few measurement inputs are available to estimate a large number of unknown state variables (in contrast to transmission systems, where the higher measurement coverage leads to over-determined systems). Even if meter placement techniques tailored to DSs have been proposed (see for example [4], [5]), it is unlikely to foresee a thorough coverage of the grid with meters in short times. Therefore, specific SE approaches have to be conceived, which must work relying on very sparse measurement information.

In the literature, pseudo-measurements are often used to deal with this problem. Pseudo-measurements may be created to model the power consumption or generation at the buses using statistical information accessible to the DSO. Their availability (at each non-measured load or generation bus) allows to solve the problem of observability and to obtain a slightly over-determined system that can be processed with conventional SE techniques, like the Weighted Least Squares (WLS) method. As an example, [6] shows how to use standard profiles of different customers' categories to create pseudo-measurements. Nevertheless, such approach may not be always easy to implement in practice, above all when considering new categories of loads (for which no statistical information is yet available) or DG (which highly depends on the weather conditions). Alternative approaches for creating pseudo-measurements via Artificial Neural Networks (ANNs) are proposed in [7] and [8]. These require the design of customized ANN models for each grid and they carry some of the drawbacks associated to neural networks, such as dependency on a black box model and need for re-training at every change in the grid or measurement configuration.

As a matter of fact, creating pseudo-measurements is not always feasible or it may be a tedious task for which DSOs do not have well-established procedures in place. For this reason, alternative SE techniques able to deal with the low number of measurement inputs present in distribution grids, without requiring the definition of additional pseudo-measurements, have been recently proposed. Some of these proposals rely on the use of ANNs [9]–[15]. The underlying idea is to train an ANN via a large set of power flow simulations that samples the space of possible operating conditions of the grid. The trained ANN can then estimate the operating state of

M. Pau is with the Department of Grid Planning and Operation of the Fraunhofer Institute for Energy Economics and Energy System Technology, Kassel, Germany (email: marco.pau@iee.fraunhofer.de).

P. A. Pegoraro is with the Department of Electrical and Electronic Engineering of the University of Cagliari, Cagliari, Italy (email: paolo.pegoraro@unica.it).

Dr. Pegoraro's work was partially funded by Fondazione di Sardegna for the research project "IQSS", Information Quality aware and Secure Sensor networks for smart cities".

the grid using just the few available measurements in input [9], [10]. A problem related to such approach is again that the model has to be updated for any change in the grid, which may be time-demanding and not suitable for real-time operation. In [11], two deep neural networks are used for SE and topology identification, and transfer learning is adopted to try to speed-up the re-training of the ANN model when the topology changes. In [12], the status of the switches is directly embedded in the ANN model, which however leads to a more complex training and design of the neural network. Another issue is the lack of transparency in the operation of the ANN. In [13], Graph Neural Networks are adopted to take into account the structure of the grid model and physical constraints are added to comply with the power flow equations. Other attempts to integrate power system physical laws in the ANN design have been proposed in [14] and [15], but in general ANN-based estimators still lack explainability which may bring some implications in terms of reproducibility and, sometimes, of trustworthiness.

Close to data driven approaches, some model-based solutions have been also proposed. The use of tracking estimators based on Kalman filters or regularized least squares models is one of the options [16], [17]. Here predictions or previous SE results can help reaching the observability, but a starting SE solution is usually needed to initialize the model and arbitrary assumptions may exist for the modelling of the grid dynamics. Another option gaining popularity is the use of matrix completion methods, whose aim is to determine missing values in low-rank matrices. SE proposals relying on this approach can be found in [18], [19]. The main issue with this technique is that it may still need a relatively high number of measurements to work well and that it involves a complex optimization process, which may be not always suitable for large grids and real-time operation. In [20], an optimization based on an interior-point solver is built for SE purposes, but a priori information on the load and generation values is still used to define inequality constraints for the power injections.

This paper offers an alternative solution for the monitoring of unobservable distribution grids, which has the advantage to be simple and to rely on the physical measurement model and the well-known WLS theory. A first idea was presented in [21], but there a simplified SE model was used, which was not able to deal with the heterogeneity of customers present in typical DS scenarios. In this paper, the WLS-model has changed and improved to allow a much more generalized application of the method, taking into account that different types of loads and generation may coexist behind each node of the grid. The main contribution of this paper thus concerns the design of a new WLS estimator for distribution grids able to work with very few measurements in grid scenarios typically considered as unobservable. To this purpose, a new SE idea is introduced, where the concepts of allocation factors and WLS are merged leading to an innovative approach for performing SE in unobservable distribution grids. It is worth noting that the idea of using allocation factors has been adopted also in [22], but in that proposal they are rather used to tune the standard load profiles for creating pseudo-measurements, which, conceptually, is a completely different solution from the method that will be illustrated here. In the following, the proposed SE method is described by presenting:

- The mathematical details of the newly proposed SE formulation, underlining the modifications with respect to the classical WLS model and thereby guaranteeing the replicability of the proposed method.
- The minimum requirements to ensure the applicability of the proposed estimator, considering both equivalent single-phase and three-phase unbalanced grids.
- A critical analysis of the performance of the SE method, using different settings and measurement configurations, proving the strengths of the proposed approach.

The remainder of the paper is organized as follows. Section II provides background information about SE. Section III dives into the idea and the mathematical formulation behind the proposed SE solution and discusses the minimum requirements to successfully employ the proposed SE method. Section IV validates it via tests carried out in different scenarios and conditions. Finally, Section V summarizes the achievements of this work and concludes the paper.

II. Distribution System State Estimation

This Section provides some background information about the two Distribution System State Estimation (DSSE) concepts at the basis of the proposed SE formulation, namely the WLS and the load allocation method.

A. Weighted Least Squares Estimation

The WLS method is a well-known technique for estimating unknown variables given a set of inputs that leads to an over-determined system of equations. In the power system context, bus voltages are often adopted as state variables of the system, while the field measurements are the inputs [23]. Despite being a relatively old method, the WLS is still the most used technique to perform SE in the control centers of grid operators. The main reasons are its accuracy (under Gaussian conditions, the WLS is a maximum likelihood estimator), its computational efficiency (the WLS solution can be found algebraically without the need of sophisticated optimization procedures), its simplicity and explainability, which allow understanding how measurements contribute to the final estimation results [24] and deriving the uncertainty characteristics of the SE output.

The WLS relies on the following measurement model:

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

where $\mathbf{z}$ is the vector of measurement inputs, $\mathbf{x}$ is the vector of state variables, $\mathbf{h}(\cdot)$ is the vector of measurement functions relating the input measurements to the state variables in $\mathbf{x}$, and $\mathbf{e}$ is the vector of errors in the input measurements, which is considered a zero-mean random vector in the following.

Given such measurement model, the objective of the WLS method is to minimize the following cost function:

$$f_{\text{obj,WLS}}(\mathbf{x}) = [\mathbf{z} - \mathbf{h}(\mathbf{x})]^T \cdot \mathbf{W} \cdot [\mathbf{z} - \mathbf{h}(\mathbf{x})] \tag{2}$$

where $\mathbf{W}$ is a weighting matrix chosen as the inverse of the covariance matrix of the measurement inputs and ^{T} indicates the transpose operator. In (2), the goal is to minimize the Mahalanobis distance between measurements and estimated measurements exploiting the existing measurement redundancy. To this purpose, the weights in $\mathbf{W}$ are tuned to give larger importance to the more accurate measurements, so that their better accuracy is reflected into the estimation process and they have more influence on the SE results.

The minimum of the cost function in (2) can be found forcing its gradient to be equal to zero. Since the measurement functions are typically non-linear, an iterative Gauss-Newton process is needed, which eventually leads to solving the following equation system at each generic iteration k [23]:

$$\mathbf{G} \cdot \mathbf{\Delta x} = \mathbf{H}^T \cdot \mathbf{W} \cdot [\mathbf{z} - \mathbf{h}(\mathbf{x}_{k-1})] \tag{3}$$

In (3), $\mathbf{H}$ is the Jacobian of the measurement functions in $\mathbf{h}$, $\mathbf{G} = \mathbf{H}^T\mathbf{W}\mathbf{H}$ is a Gain matrix, and $\mathbf{\Delta x} = \mathbf{x}_k - \mathbf{x}_{k-1}$ is the vector used to update the estimated state variables $\mathbf{x}$ between the generic iterations $k-1$ and k. This iterative process is carried out until a certain convergence criterion is satisfied. Usually, this is given by the maximum change (in absolute value) of the state variables, which must fall below a given threshold ϵ, namely $\|\mathbf{\Delta x}\|_\infty < \epsilon$

In DSs, several WLS-based DSSE formulations adapted to the peculiarities of these grids have been proposed. These include, among others, the three-phase modelling of the electrical grid (e.g., [25], [26]), the use of different state variables like branch currents (e.g., [27], [28]) and, above all, the integration of pseudo-measurements to achieve system observability [29], which is essential for the application of the WLS-based formulation in its original form.

B. Load Allocation Concept

Until some decades ago, distribution grids were mostly operated as passive networks and hence did not require a particularly complex management. From a monitoring perspective, often a measurement point was available only at the main substation and the operating conditions of the entire grid were estimated starting from it. As discussed in [30], a common approach was to estimate a load allocation factor K_{load} as follows:

$$K_{load} = \frac{S_{meas}}{\sum_{i=1}^{N} \bar{S}_i} \tag{4}$$

where S_{meas} is the apparent power measured at the main substation, $\bar{S}_i$ is the rated power of the i-th bus, and N is the total number of buses in the grid. Such allocation factor basically is the per unit level of power (with respect to the rated one) assigned to each load depending on the overall power seen at the main substation. Following this idea, the power consumption at the i-th bus could be computed as:

$$S_i = K_{load} \cdot \bar{S}_i \tag{5}$$

Typical power factors could be used to split the apparent power in its active and reactive power components. Having the estimated power at each node, and using the voltage measurement at the main substation, it was then possible to derive the complete voltage profile of the grid performing a standard power flow calculation.

Such an approach is clearly not suitable for modern DSs, as the presence of DG prevents assuming that all nodes of the grid behave in a similar way (i.e., that they work at the same per unit level). Nevertheless, the idea behind the use of allocation factors is adapted in this paper to fit a scenario where heterogeneous loads and DG are connected to the distribution grid. As shown in the next Section, combining allocation factors with the WLS method eventually allows the monitoring of unobservable distribution grids.

III. Proposed DSSE Model

A. General Concept

The underlying idea behind the load allocation method presented in Section II was that a single power measurement may be used (together with the voltage at the main substation) to derive a rough estimation of the operating conditions of the entire grid. The only unknown in such a problem was the load allocation factor and, hence, a single power measurement was sufficient to determine its value.

In the current DS scenario, heterogeneous types of loads and DG are present. However, it is still possible to imagine each bus of the grid as the aggregation of a limited number of load and generation clusters, each one associated with a specific allocation factor. As an example, let us assume that the i-th bus of the grid is a medium to low voltage (MV/LV) substation and that the mix of downstream users connected to the LV grid includes Q_i different clusters of loads and/or generators. Generalizing (5), the apparent power injection at this node can be expressed as:

$$S_i = \sum_{h=1}^{Q_i} K_h \cdot \bar{S}_{i|h} \tag{6}$$

where K_h is the allocation factor associated with the h-th cluster and $\bar{S}_{i|h}$ is the total rated power of the customers connected downstream the considered node i and belonging to such a cluster h.

Inspired by the load allocation method, the key assumption in the proposed DSSE algorithm is that different buses of the grid share the same allocation factors when referring to the same type of load (or generation) cluster. This is equivalent to say that the power consumed (or generated) within each cluster would be always at the same per unit level (with respect to the rated one) for each node of the grid. This is essentially the same assumption existing in the load allocation method, which is here extended to consider the possible presence of multiple clusters or categories of loads and generators. In case of clusters associated to DG based on renewable energy sources (for example, PV or wind plants), this assumption sounds realistic, as it is reasonable to assume that plants scattered over a relatively small geographical area are subject to similar weather conditions and, consequently, that they may have a quite similar generation behavior. In the case of loads, different clusters can be created to aggregate loads with statistically similar consumption patterns. This is similar to what done when tuning pseudo-measurement according to different standard load profiles, considering the category of the end-users [6]. Conceptually similar assumptions have been already used also in other DSSE proposals, like [31] and [32], where different types of loads or generators were considered to have very high correlations.

Using the above assumption, the DSSE problem can eventually be seen as the estimation of a limited number of allocation factors associated to different clusters of loads and generators. In this way, a small number of measurements would be sufficient to estimate these allocation factors. In the proposed DSSE algorithm, this idea is integrated into a WLS formulation, which gives the possibility to estimate simultaneously the discussed allocation factors as well as the bus voltages.

B. Proposed WLS Formulation

As mentioned above, the proposed WLS aims at estimating simultaneously bus voltages and allocation factors. To this purpose, the state vector $\mathbf{x}$ is augmented to include both allocation factors and voltage variables as follows:

$$\mathbf{x} = [\mathbf{V}_{r,A}^T, \mathbf{V}_{x,A}^T, \mathbf{V}_{r,B}^T, \mathbf{V}_{x,B}^T, \mathbf{V}_{r,C}^T, \mathbf{V}_{x,C}^T, K_1, \cdots, K_Q]^T \tag{7}$$

where $\mathbf{V}_{r,\phi}$ and $\mathbf{V}_{x,\phi}$ are the vectors of the real and imaginary voltages at phase ϕ (with $\phi \in \{A, B, C\}$), K_h is the allocation factor of the h-th cluster, and Q is the total number of clusters identified in the grid. Note that the proposed state vector refers to a three-phase representation, but it can be easily adapted to the single-phase case. Also, as shown in [28], note that the imaginary voltage of a reference bus can be removed from the state vector if no Phasor Measurement Units (PMUs) or, more in general, no absolute phase-angle measurements are available in the input measurement set.

The inputs considered in the measurement vector $\mathbf{z}$ include: i) the measurements collected from the field, for which the same measurement functions as in the conventional WLS estimators can be used (see the Appendix for more details); ii) zero values, which result from the following balance equations defined for the current injection of each bus of the grid (including also zero-injections) and each phase of the system:

$$\begin{aligned} -\sum_{h\in\Pi_i} K_h \cdot \bar{I}_{ri,\phi|h} + \sum_{\psi\in\Psi}\sum_{k\in\Gamma_i} [g_{ik,\phi\psi}(V_{ri,\psi} - V_{rk,\psi}) \\ - b_{ik,\phi\psi}(V_{xi,\psi} - V_{xk,\psi})] = 0 \end{aligned} \tag{8}$$

$$\begin{aligned} -\sum_{h\in\Pi_i} K_h \cdot \bar{I}_{xi,\phi|h} + \sum_{\psi\in\Psi}\sum_{k\in\Gamma} [b_{ik,\phi\psi}(V_{ri,\psi} - V_{rk,\psi}) \\ + g_{ik,\phi\psi}(V_{xi,\psi} - V_{xk,\psi})] = 0 \end{aligned} \tag{9}$$

where Π_i is the set of indexes of the clusters composing the i-th node power mix (and $|\Pi_i| = Q_i$) and $\Psi = \{A, B, C\}$; Γ_i is the set of nodes adjacent to node i; $g_{ik,\phi\psi}$ and $b_{ik,\phi\psi}$ are the real and imaginary components of the series admittance of the branch between nodes i and k (self admittance if $\psi = \phi$ or mutual admittance between different phases if $\psi \neq \phi$); $V_{ri,\psi}$ and $V_{xi,\psi}$ are the real and imaginary voltage at bus i and phase ψ; $\bar{I}_{ri,\phi|h}$ and $\bar{I}_{xi,\phi|h}$ are the real and imaginary components of the rated current for the h-th cluster available at node i, which are obtained as follows:

$$\bar{I}_{ri,\phi|h} + j\bar{I}_{xi,\phi|h} = \frac{\bar{P}_{i,\phi|h} - j\bar{Q}_{i,\phi|h}}{V_{ri,\phi} - jV_{xi,\phi}} \tag{10}$$

In (10), the considered voltages are temporary estimates obtained during the WLS procedure, while $\bar{P}_{i,\phi|h}$ and $\bar{Q}_{i,\phi|h}$ are, respectively, the sum of the rated active and reactive powers for the users belonging to cluster h connected behind bus i at phase ϕ. The active and reactive components of rated power can be extracted from the apparent power using typical power factors associated to the considered category of load or generation.

The relationships in (8) and (9) essentially give constraints to the power injection computed through the voltage variables which should be equal to the power injection given in (6). Here, it is worth noting that the associated zero values inserted in the measurement vector $\mathbf{z}$ are the result of such balance equations and, thus, they should not be misinterpreted as representing in general an assumption of zero current injection. For load/generation buses, the equality constraint assumed in the balance equations (8) and (9) will not be perfect as, in practice, the assumption that all loads (or generators) of the grid belonging to the same cluster behave in the same way (namely, that they work at the same per unit level of power with respect to their rated value) is just a statistical approximation. However, (8) and (9) become loose constraints when used as equivalent measurements in the vector $\mathbf{z}$. Indeed, the lack of exactness in the assumption behind each of these relationships can be modelled as the uncertainty of the corresponding equivalent measurement and, therefore, be mapped into the WLS through the weighting matrix. The ideal procedure to define the needed weights would first require a statistical characterization of the variability of loads (or generators) belonging to the same cluster. This statistical information could be used to describe the a-priori uncertainty associated with the allocation factors in terms of their variance $\sigma^2_{K_h}$. The law of propagation of the uncertainty [33] can be then applied to (8) and (9) to find the resulting variances $\sigma^2_{\mathrm{eq}_{ri,\phi}}$ and $\sigma^2_{\mathrm{eq}_{xi,\phi}}$ for the equivalent measurements expressing the balance of real and imaginary current injection at node i and phase ϕ. Neglecting in first approximation the uncertainty contributions associated with the power factor used to convert the apparent power into its active and reactive components, and possible correlations between allocation factors[1], the following holds:

$$\sigma^2_{\mathrm{eq}_{ri,\phi}} = \sum_{h\in\Pi_i} \bar{I}^2_{ri,\phi|h} \cdot \sigma^2_{K_h} \tag{11}$$

$$\sigma^2_{\mathrm{eq}_{xi,\phi}} = \sum_{h\in\Pi_i} \bar{I}^2_{xi,\phi|h} \cdot \sigma^2_{K_h} \tag{12}$$

As done also for the actual measurements, the weights introduced in the weighting matrix must be then chosen as the inverse of the variances in (11) and (12). Since in practical scenarios it may be difficult to have accurate definitions of the a priori uncertainties associated to the allocation factors,

[1]This assumption is realistic since the prior variability is mainly tied to the cluster type.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2024.3387498

in Section IV, some tests are presented to show the impact of an inaccurate choice of such weights.

Overall, it can be noted that, in the proposed formulation, the use of pseudo-measurements can be completely avoided as the additional equations needed to make the WLS model over-determined are given by the balance equations (8) and (9). The definition of such balance equations requires only the knowledge of the rated powers installed behind each node, without any a priori forecast or guess of the allocation factors or of the power injections. In addition, it is also worth noting that the balance equations and the associated uncertainties hold also for the case of zero injections (ZIs). In fact, in case of ZIs, the terms in (8) and (9) associated with the allocations factors would disappear (due to the null value of the rated currents associated to each customer cluster) thus leading to the classic ZI equation. Similarly, in (11) and (12), the resulting variances would be equal to zero, as it is appropriate to expect, given that no uncertainty exists about the knowledge of the ZI. In the algorithm tested in the following, the weight for the ZIs is simply replaced with a large but limited value, keeping ill-conditioning under control. It is however worth mentioning that, to further improve the conditioning of the WLS algorithm, the ZIs could also be modelled as constraints and integrated in the WLS via a Lagrangian approach, which was not needed in the tests reported in Section IV.

To complete the design of the proposed WLS estimator, the partial derivatives appearing in the Jacobian matrix must be defined. Regarding the derivatives with respect to the voltage variables, it is worth noting that no modifications are required with respect to the original version of the WLS. This holds also true for the equivalent measurements associated with the current injection balance, which keep exactly the same derivatives used in the original WLS for the pseudo-measurements of power injections (which, when using a formulation with rectangular voltages as state variables, are converted into equivalent currents). Regarding the derivatives with respect to the allocation factor state variables, the only derivatives differing from zero are those associated with the equivalent measurements associated with the current injection balance. As it can be easily derived from (8) and (9), these derivatives are:

$$\frac{\partial h_{\mathrm{eq}_{ri,\phi}}}{\partial K_h} = -\bar{I}_{ri,\phi|h} \qquad \frac{\partial h_{\mathrm{eq}_{xi,\phi}}}{\partial K_h} = -\bar{I}_{xi,\phi|h} \tag{13}$$

thus showing that the only additional non-zero entries in this Jacobian sub-matrix are those associated with the clusters actually present in the power mix of node i.

C. Considerations on Observability

Starting from the assumption that the monitoring stations are few and thus the base system of equations linked to real measurements is under-determined, observability needs to be guaranteed by the above-described equivalent measurements derived from the balance equations in (8) and (9). First of all, like in the WLS with pseudo-measurements, at least one voltage measurement is needed also in the proposed DSSE. Since the equivalent measurements are defined for $N-1$ buses, the combination of a voltage measurement point with these equivalent measurements allows having as many measurements as the number of voltage state variables. The number of additional state variables linked to the allocation factors then determines the number of other real measurements needed in the field. With $Q=1$, i.e. a single allocation factor, any additional measurement guarantees observability. With more allocation factors involved, Q additional measurements providing further information about the voltage profile and/or power flows are necessary (and sufficient).

In addition to the requirement above, since additional unknowns associated with the allocation factors are present in the proposed DSSE model, it is important also to analyze the corresponding columns in the augmented Jacobian matrix $\mathbf{H}$. In particular, the following condition must hold true:

$$\operatorname{rank} \mathbf{H}_K = Q \tag{14}$$

where $\mathbf{H}_K$ represents the submatrix of the Jacobian composed of the columns associated with the derivatives with respect to the allocation factors K_h. Since all the derivatives of the real measurement functions with respect to K_h are null because, in the proposed formulation, there is no dependency on allocation factors, the submatrix of $\mathbf{H}_K$ associated with them is a zero matrix, and thus the analysis of the rank can be limited to the rows corresponding to equivalent measurements. More specifically, the condition in (14) indicates that the power mix of the network buses needs to be linearly independent. For instance, having exactly the same mix for all the buses in terms of both involved clusters and rated powers would lead to linear dependence among the columns of $\mathbf{H}_K$ and, thus, to indistinguishable allocation factors. The criterion in (14) is straightforwardly met when $Q = 1$ and, generally, it is easily met also when $Q > 1$ if the nature of the node powers is composite (i.e., if the nodes have different rated powers associated to the different clusters).

Finally, it is interesting to note that that the conditions above lead to different measurement installation requirements in balanced (single-phase) and unbalanced (three-phase) grids. In particular, in the three-phase scenario observability may be met more easily, because each measurement point provides three independent measurement values (one for each phase) that would allow estimating up to three different allocation factors. In the equivalent single-phase scenario, instead, only a single measurement value is available per measurement point (the positive sequence), which would then allow estimating only a single cluster. Consequently, in an unbalanced scenario, the number of installed measurement points required to achieve the observability can be reduced by a factor of 3.

IV. Tests and Results

A. Simulation Set-up

The proposed WLS method for SE in unobservable distribution grids has been tested in different scenarios and with different measurement configurations using the three-phase unbalanced IEEE 123-bus grid shown in Fig. 1. The grid has been modified to introduce four categories of connected customers: residential and commercial loads, and PV and wind generation. In particular, each bus of the grid that has a

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2024.3387498

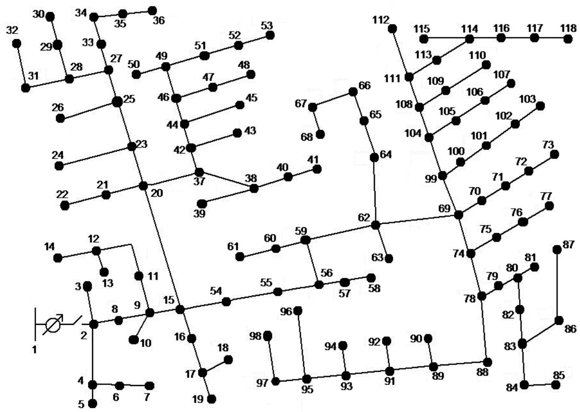

Fig. 1. IEEE 123-bus test distribution grid.

load in the default data [34] has been replaced with a mix of connected loads and generators, each one with different rated powers for each node. In general, different clusters are connected to each bus, with most of the nodes subtending at least two clusters of customers and several having all four clusters simultaneously. Overall, the grid has a total connected power of around 3.8 MW for residential loads, 2.9 MW for commercial loads, and 3.9 MW and 2.2 MW for PV and wind generation, respectively. The used network parameters are the same given in [34], whereas voltage regulators and a few buses connected to open switches have been omitted for the sake of simplicity.

Given the considered clusters, to create the loading and generation scenario of the grid, while emulating the different behavior of customers belonging to the same cluster, the power at the different buses has been extracted randomly (with uniform distribution) according to the following assumptions:

- residential loads: between $50\,\%$ and $80\,\%$ of rated power;
- commercial loads: between $30\,\%$ and $60\,\%$ of rated power;
- PV plants: between $30\,\%$ and $40\,\%$ of rated power;
- wind plants: between $20\,\%$ and $40\,\%$ of rated power;

The power at each node hence results from the rated power available at that node for every connected cluster and from the above-mentioned random extractions. The overall power consumption/injection of the node is eventually obtained as the sum of the power contributions of each cluster. From the resulting load and generation scenario, a power flow calculation is used to obtain the reference operating conditions of the grid that are assumed as "true". For the DSSE, measurements are then extracted from the associated reference values considering $1\,\%$ and $2\,\%$ expanded uncertainty for the voltage and power measurements, respectively, with Gaussian uncertainty distribution (coverage factor 3). To analyze the DSSE performance in different scenarios, the following measurement configurations are considered:

- *Case 0*: voltage measurements at nodes 1 and 69 and active and reactive power at branches 1-2 and 62-69.
- *Case 1*: full metering points at nodes 1 and 69.
- *Case 2*: full metering points at nodes 1, 15, 27 and 69.
- *Case 3*: full metering points at nodes 1, 15, 27, 37, 69, 78 and 111.

Case 0 is used to prove the operation of the proposed algorithm with a minimal deployment of measurements needed to reach the observability. The other cases instead allow analysing the enhancement of the accuracy performance with a slightly larger measurement redundancy. In this regard, full metering points indicate the presence of a voltage measurement at the bus, and of active and reactive power measurements in all the branches converging to the bus.

All simulation results shown in the following are obtained via Monte Carlo (MC) simulations, using $N_{\mathrm{MC}} = 5000$ MC trials, in order to have statistically meaningful results. Each MC iteration involves the random extraction of both the operating points of loads and generators (within the intervals indicated above for each cluster) and of the measurements (within their measurement uncertainty interval) from the corresponding reference values obtained via the power flow calculation. Moreover, different initialization settings for the state variables have been tested in a preliminary phase. In particular, a flat voltage was used to initialize the voltage variables, while the allocation factors have been initialized with different values between 0 and 1 (e.g., including all allocation factors equal to 0 or to 1). In all performed tests, the estimator converged exactly to the same SE results regardless of the particular values chosen for the initialization of the allocation factors, thus proving that no a-priori guess of these factors is needed and that no strict rules apply for their initialization.

For the following test results, the performance index is computed as:

$$u_{\nu} = \sqrt{\frac{1}{N_{\mathrm{MC}}} \sum_{n=1}^{N_{\mathrm{MC}}} (\hat{\nu}_n - \nu_n)^2} \tag{15}$$

where $\hat{\nu}_n$ indicates the estimated quantity (voltage magnitude or branch active power of a specific node or branch, respectively) in the n-th MC trial, and ν_n is its reference counterpart. u_{ν} is the root mean square error (RMSE), but, as will be discussed in Section IV-B, it represents also the standard deviation of the obtained zero mean estimation error. For this reason, it significantly describes the uncertainty of the algorithm. Other indices can be used, e.g., the mean absolute estimation error, but they are not reported here for the sake of brevity, since the conclusions presented in the comments to the tests below would be still the same.

B. Tests with Different Measurement Configurations

First tests have been performed to prove the possibility to run the proposed DSSE with a minimum number of measurements, thus referring to the Case 0 scenario. In such a measurement configuration, only two branches of the grid are monitored by associated power meters. Exploiting the unbalanced conditions, resulting from the different loads and generation present in the three phases of the grid, six independent power measurements are eventually available, which thus allow estimating the allocation factors of all the four clusters assumed in the grid.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2024.3387498

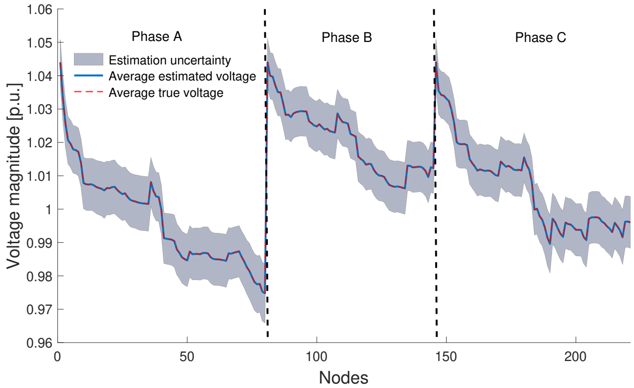

Fig. 2. Voltage magnitude estimation with minimum measurement configuration (Case 0).

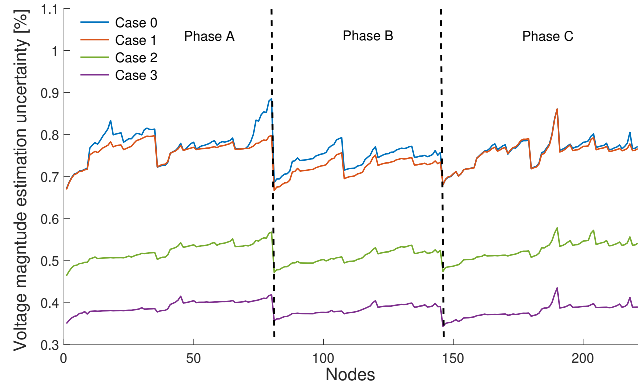

Fig. 4. Expanded uncertainty of the voltage magnitude estimation for different measurement configurations.

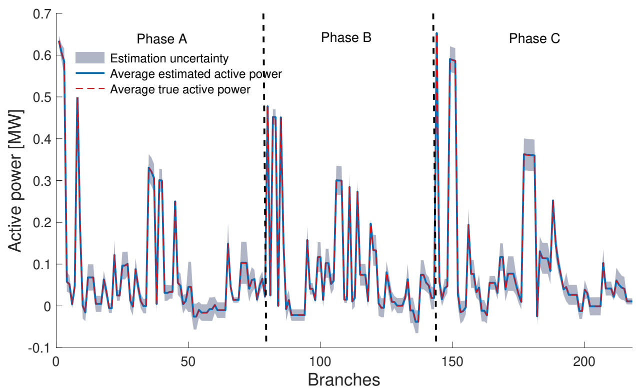

Fig. 3. Branch active power estimation with minimum measurement configuration (Case 0).

Figure 2 shows the voltage magnitude profile of the grid averaged over the MC trials, together with the expanded uncertainty (obtained from u_ν using a coverage factor equal to 3) of the DSSE results[2]. The average estimated voltage magnitude profile is reported only to confirm that, also with this new formulation, the WLS keeps its unbiasedness and is capable of correctly tracking the voltage variations along the grid. Indeed, such profile is equal to the (average) true one, as expected. More important, with this minimal measurement configuration, the expanded uncertainty of the voltage magnitude estimation ranges between $0.67\,\%$ and $0.89\,\%$, which is an outstanding result considering the very low number of measurements in the grid.

Figure 3 shows the analogous results for the active power estimation over the network branches. Also in this case, the DSSE results are unbiased, namely the estimated and true power profile averaged over the MC trials are the same. The expanded uncertainty (coverage factor equal to 3) of the estimated active power may highly vary for different branches, but it is in average (when averaging over all the network branches) equal to 24 kW and the highest peak is equal to 60 kW. As it can be inferred from Fig. 3, for some branches of the grid this can result in a very high percentage uncertainty, as some of the branch powers are very close to 0 due to the combined effects of load and generation profiles. In general, however, these levels of uncertainty are much lower than the largest powers flowing in the grid, which are an order of magnitude larger. Such results show that also with a minimum measurement configuration, the proposed estimator is able to track quite well the power flows in the grid thanks to the estimation of the introduced allocation factors.

The performance of the proposed estimator may significantly improve with a larger deployment of measurements. Figure 4 shows the enhancements achievable for the estimation of the voltage magnitude profile. Even if Case 1 has the same number of voltage measurements as in Case 0, it is possible to observe that the additional power measurements allow a better estimation of the allocation factors, which is eventually reflected also in the accuracy of the voltage estimation. Case 2 and Case 3 involve instead a larger number of voltage measurements and this allows to clearly reduce the plateau of voltage estimation uncertainty to around $0.5\,\%$ and $0.38\,\%$, respectively. It is also worth noting that these levels of expanded voltage uncertainty are compliant with the rule of thumb derived in [24], namely they are in the order of $^{U_{V_m}}/\sqrt{N_{V_m}}$, where U_{V_m} is the expanded uncertainty of the voltage meters (in these tests $1\,\%$) and N_{V_m} is the number of available voltage meters.

Figures 5 and 6 show one of the unique features of the proposed DSSE formulation, namely the capability to identify the operating point of different load or generation clusters (according to the scenario depicted in Section IV-A, the average allocation factors are 0.65 and 0.45 for the residential and commercial loads, and 0.35 and 0.30 for the PV and wind generation, respectively). In the figures, the boxes represent the first and third quartile of the allocation factor estimates, the line inside them is the median and the whiskers are the maximum and minimum estimates that are identified as non-outliers. The outliers are those estimated values beyond $150\,\%$ of the interquartile range. In Case 0, since only two

[2]From here on, and differently from Fig. 1, the node index in the figures is a sequential number for all the three system phases starting from bus 1 of phase A to the last bus of phase C. Indeed, in the network there are three-, two- and single-phase buses.

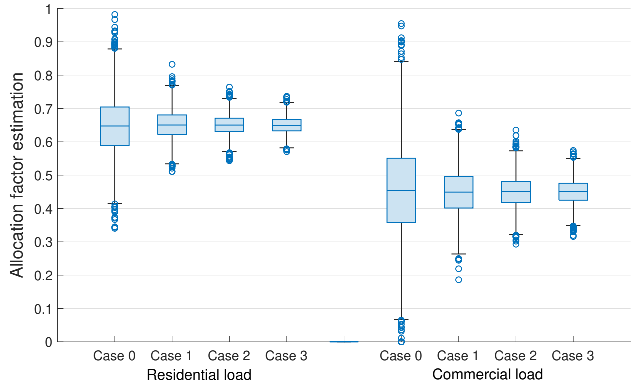

Fig. 5. Statistics of the estimated load allocation factors for different measurement configurations.

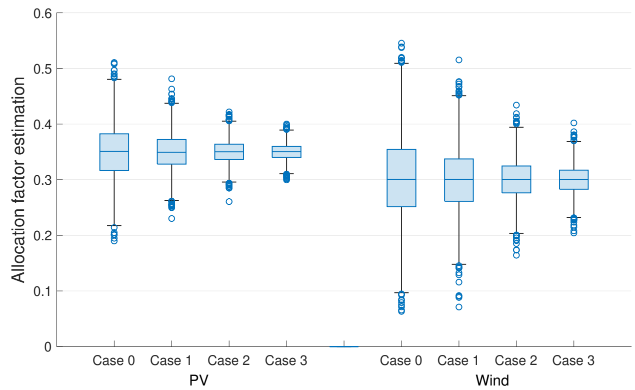

Fig. 6. Statistics of the estimated generation allocation factors for different measurement configurations.

power measurements are available, the estimation of the allocation factors is not always accurate, as the loads and generators may be combined in multiple ways (with different allocation factors) to give results consistent with the available measurements. However, already with Case 1, it is possible to observe that the allocation factors are estimated much more accurately (e.g., 0.65 ± 0.12 for the residential loads and 0.35 ± 0.09 for the PV generation). This improvement continues till Case 3 where, even if only 7 out of the 118 active nodes of the grid are monitored, the proposed formulation allows estimating accurately all the allocation factors (e.g., with an uncertainty of only ± 0.04 for the PV cluster). In general, the different performance for different clusters are associated with the number of loads/generators belonging to each cluster. In the created scenario, for example, most of the nodes have residential loads and PV generation connected behind them, while less nodes have commercial loads or wind generation. This explains why the average behavior of these last clusters is captured slightly less accurately.

Finally, Table I shows some statistics about the convergence properties of the proposed DSSE for the MC simulations discussed above. As it can be seen, the algorithm typically converges in only five iteration, except for Case 0, where the low measurement redundancy can sometimes lead to six

TABLE I
AVERAGE NUMBER OF ITERATIONS AND COMPUTATION TIME

Case scenario	Avg number of iterations	Avg execution time [ms]
Case 0	5.51	84.27
Case 1	5.00	83.49
Case 2	5.00	91.93
Case 3	5.00	109.93

TABLE II
DIFFERENT OPERATING CONDITIONS SCENARIOS: VARIABILITY RANGES FOR ALLOCATION FACTORS

Customer cluster	Allocation range	
	Scenario A	Scenario B
Residential load	20% - 40%	20% - 40%
Commercial load	30% - 60%	10% - 40%
PV generation	70% - 90%	0%
Wind generation	20% - 30%	20% - 80%

iterations. In all the considered scenarios, the execution time is in the order of hundred milliseconds (the algorithm was implemented in MATLAB and run on a 1.90 GHz processor with 16 GB of RAM).

C. Tests with Different Operating Conditions

Previous tests considered different operating conditions, but using predefined ranges for the extraction of the powers of loads and generators within each cluster. In this section, the percentage range of each cluster is modified to prove that the considerations drawn in Section IV-B still hold also when looking at different operating scenarios. In particular, two further scenarios are presented, whose load/generation percentage range (with uniform distribution) for each cluster is given in Table II. Scenario A may represent a time step in the afternoon, where a high generation from PV plants and a relatively low consumption from residential loads occur. Scenario B, instead, could be representative of a time step in the night with no PV generation, a low consumption from residential and commercial loads, and a highly variable generation from wind plants. All the results discussed in this section are created using Case 1 measurement configuration.

Figure 7 shows the expanded uncertainty (coverage factor equal to 3) for the voltage magnitude estimation for Scenarios A and B in comparison with the scenario of Section IV-B (default scenario, using Case 1 too). Clearly, slightly different profiles of uncertainty exist due to the different operating conditions under analysis. However, it is possible to observe that, also in these new scenarios, the voltage magnitude estimation uncertainty ranges for most of the nodes between $0.7\,\%$ and $0.8\,\%$, which is in line with results obtained for the default scenario and coherent with the levels of uncertainty expected according to the adopted measurement configuration [24]. In a similar way, also the allocation factors are estimated consistently with what shown in Figs. 5 and 6, namely the

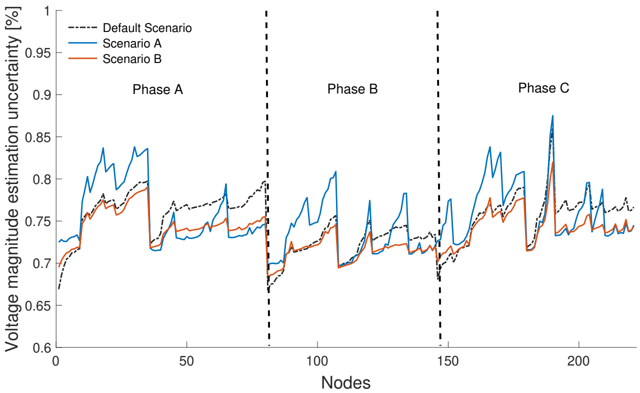

Fig. 7. Expanded uncertainty of the voltage magnitude estimation for different operating scenarios.

estimates are very close to the average value of the considered ranges, with slightly more accurate estimations for the residential load and PV clusters. A further aspect should be highlighted concerning the estimation of the allocation factor associated with the PV cluster in Scenario B. The presented results refer to the case in which this allocation factor is still treated as a completely unknown state variable to prove the robustness of the proposed algorithm. It is however trivial to notice that, if an allocation factor is known a priori (like for the PV generation during the night), then this information can be directly embedded in the algorithm and such allocation factor can be simply removed from the set of state variables.

Similar tests have been run considering also other ranges of variability for the allocation factors, here including also scenarios with very high variability (e.g., $10\,\%$–$80\,\%$ for the residential loads). In general, a larger variability of the power profiles within the same cluster may slightly degrade the accuracy performance, both in terms of voltage and allocation factor estimation. However, obtained results, considerations and capabilities of the estimator still remain very similar to those discussed for the previous test cases. For this reason, and for the sake of brevity, a detailed presentation of the results for these additional test scenarios is here omitted.

D. Impact of State Estimation Weights

One of the peculiarities of the proposed DSSE algorithm is the use of the equivalent measurements corresponding to current injection balance defined in (8) and (9) as a function of the individual contributions from each of the connected clusters of customers. While all the other settings of the DSSE algorithm are quite straightforward, the definition of the weights for these equivalent current injections may be more troublesome, as it should reflect the variability resulting due to the different behavior of customers belonging to the same cluster. Equations (11) and (12) assume the knowledge of the variability of the allocation factors, but in practical scenarios a precise information on such variability may be missing. For this reason, additional tests have been performed to assess the robustness of the proposed DSSE algorithm to an erroneous choice of the equivalent current injection weights.

TABLE III
IMPACT OF THE WEIGHTS OF THE CURRENT INJECTION EQUIVALENT MEASUREMENTS ON THE ESTIMATION OF THE ALLOCATION FACTORS

Cluster	Weighting scenario	Statistic	
		Median	5^{th}-95^{th} percentile
Residential load	Case 3	0.6497	0.6081 - 0.6912
	Case 3a	0.6497	0.6079 - 0.6912
	Case 3b	0.6495	0.6079 - 0.6919
	Case 3c	0.6474	0.6011 - 0.6961
Commercial load	Case 3	0.4514	0.3872 - 0.5110
	Case 3a	0.4513	0.3864 - 0.5122
	Case 3b	0.4521	0.3862 - 0.5127
	Case 3c	0.4483	0.3747 - 0.5208

In the previously presented results, the values of σ_{K_h} to be used in (11) and (12) for each cluster were chosen considering the maximum variability of the allocation factor (from its average value) divided by $\sqrt{3}$ (due to the assumed uniform distribution). For testing purposes, further simulations have been run with the default scenario of Section IV-B and considering three different weights configurations for the algorithm:

- *Case a*: the values of σ_{K_h} used in the algorithm are scaled up by a factor 3 with respect to the correct ones, thus assuming less accurate equivalent current injection measurements, to assess the results with a much more conservative assumption about the variability of allocation factor within each cluster (lack of a-priori information);
- *Case b*: the adopted values of σ_{K_h} are fixed and equal to ${}^{0.3}/\sqrt{3}$ for all the clusters, thus assuming a $30\,\%$ maximum variability of allocation factor in each cluster, to assess the results when the same default weight is used regardless of the presence of less variable clusters. Also in this case, uncertainty of constraints associated with equivalent measurements is thus overestimated at execution time;
- *Case c*: values of σ_{K_h} are assumed equal to ${}^{0.01}/\sqrt{3}$ in the algorithm for all the clusters, thus assuming a prior $1\,\%$ maximum variability in each cluster, to assess the results with a large mismatch in the assumption about the variability within each cluster and, in particular, a strong underestimation of the uncertainty of equivalent measurements constraints.

Performed tests show that the choice of the equivalent current injection weights has only a very small impact on the final results, if no large mismatches exist. This impact slightly grows when more measurements are used (Case 3), probably because the additional power measurements bring additional constraints for which it becomes important to correctly model all the input weights. Table III shows the statistics for the estimation of the load allocation factors in Case 3, including the three additional scenarios with the incorrect modelling of the weights. It is possible to observe that using more conservative standard deviations (Case 3a) or a wrong proportion of the weights (Case 3b) may lead to slightly larger uncertainties with respect to the base case, but, as mentioned, the differences are always extremely small. However, when gross errors exist in the definition of the weights (Case 3c), this can lead to a clearer degradation of the results (affecting also the median of the estimated allocation factors). Similar results and considerations hold also for the generation allocation factors.

Such results suggest that a rough estimation of the variability of the power consumption/generation for customers belonging to the same cluster is still sufficient to obtain meaningful DSSE results. Indeed, in the simulated scenarios, the very small differences obtained for the estimation of the allocation factors in Case 3a and 3b do not lead to any appreciable difference in the estimation of the voltages and currents of the grid with respect to the base case with correct setting of the weights. At the same time, the results of Case 3c highlight that it is still important to have a realistic, non-random definition of the weights, in order to prevent an avoidable degradation of the DSSE results.

E. Comparison with Conventional WLS

An additional series of tests has been performed to compare the proposed DSSE method (Allocation-Factor WLS, AF-WLS) with a conventional WLS based on pseudo-measurements [28]. The aim is to highlight some risks that may occur when having a rough definition of the pseudo-measurements to achieve grid observability. All the following tests adopt the measurement configuration given in Case 1.

Figure 8 compares the accuracy performance of the proposed estimator to those of a conventional WLS equipped with different pseudo-measurements. In particular, 'WLS 1' defines the pseudo-measurements via (6) and using the average value of the allocation factors for each cluster. The associated weights are computed using (11) and (12). It is possible to see that in this case the accuracy performance of the proposed estimator and the conventional WLS are very similar. However, the conventional WLS requires the a priori knowledge of the allocation factors for defining the pseudo-measurements, whereas in the proposed method this requirement does not exist as the allocation factors are actually estimated.

'WLS 2' uses the same approach of WLS 1 to create the pseudo-measurements, but in this case the knowledge of the allocation factors is assumed to be biased. A negative and a positive bias of 0.15 are considered for the load and generation clusters, respectively. The weights are again computed using (11) and (12). This case emulates a scenario where it is not possible to accurately determine the pseudo-measurements, but they are anyway used to achieve observability. The results show that in such a case an important degradation of the accuracy performance may occur, as visible above all in Phase A and Phase C (which are the most loaded phases). This risk is clearly avoided with the proposed method, as there is no a priori definition of the pseudo-measurements.

'WLS 3' uses the same pseudo-measurements created for WLS 2, but it adopts a default value of pseudo-measurement uncertainty (equal to 50 % of the pseudo-measured power consumption or generation) for the creation of the associated weights. This case reflects a scenario where, beyond the pseudo-measurement values, it is also not possible

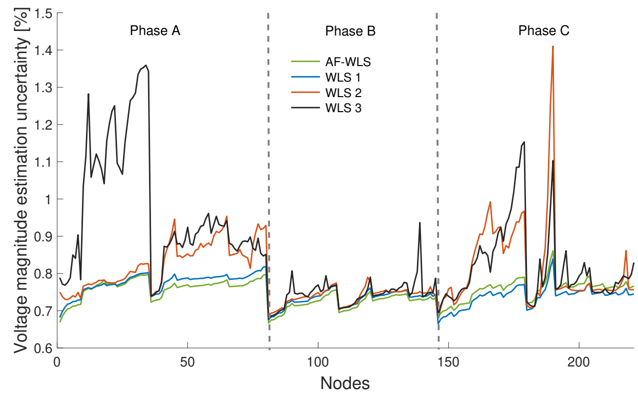

Fig. 8. Accuracy performance comparison between proposed method and conventional WLS with different choices of pseudo-measurements (Case 1).

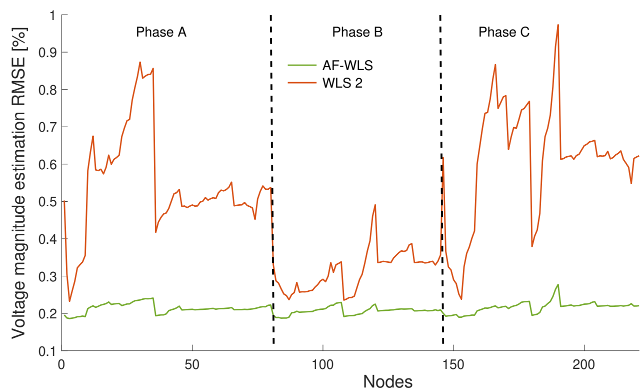

Fig. 9. Accuracy performance comparison between proposed method and conventional WLS in case of loss of power measurement at bus 1 (Case 1).

to accurately determine the uncertainty around the pseudo-measurements. The results prove that also the definition of the pseudo-measurement weights is critical, as visible above all in Phase A, where an important increase of the voltage estimation uncertainty is obtained due to the simplistic assumption behind the definition of the pseudo-measurement uncertainties.

Finally, Fig. 9 shows another particular scenario where the use of pseudo-measurements may potentially lead to misleading results. In this scenario, the measurement configuration of Case 1 is used, but it is assumed that the power measurement at bus 1 is temporarily unavailable and cannot be used as input to the state estimator. The pseudo-measurements are created with a bias, according to what previously described for WLS 2. Figure 9 depicts the accuracy performance in terms of voltage magnitude estimation RMSE, showing that the loss of the power measurement at bus 1 would severely affect the performance of the conventional WLS. The reason for this is that, with the loss of this power measurement, a large number of pseudo-measurements become critical measurements that cannot be refined during the estimation process. In this case, the bias in the pseudo-measurements is hence not compensated and it propagates to the voltage profile, eventually leading to a biased voltage estimation and to significantly higher RMSE values. It is worth noting that in the proposed estimator, instead, the estimation of the allocation factors (which is still possible thanks to the power measurements at bus 69) allows an unbiased estimation of all the power injections and, consequently, also an accurate estimation of voltage profile.

As demonstrated by the above tests, the proposed estimator allows avoiding some important issues associated to the choice of the pseudo-measurements and of their weights. This is of utmost importance in future scenarios, since the diffusion of new loads (such as electric vehicles, heat pumps, etc.) will lead to a lack of statistical information and to a much more complex definition of the pseudo-measurements.

F. Bad Data Detection Performance

One of the main functionalities for a state estimator is the capability to detect and identify possible bad data appearing in the input measurement set. Even though the low measurement redundancy at distribution level makes it hard to recognize some types of bad data, this section aims at investigating if the tools adopted for bad data detection and identification in the conventional WLS can be still successfully applied also within the proposed WLS formulation. To this purpose, the Largest Normalized Residual (LNR) approach was implemented in the DSSE algorithm and MC tests were run, applying bad data to the voltage magnitude measurements when considering the Case 2 measurement configuration (which ensures that voltage measurements do not form a critical pair and that bad data are both detectable and identifiable).

Obtained results proved that the LNR test can be successfully employed also in the proposed WLS formulation and allows identifying voltage bad data with a very good sensitivity. As an example, bad data with a negative offset as low as $2\,\%$ for the voltage measurement at bus 27 are successfully identified in $58.7\,\%$ of the cases, while bad data with a $2\,\%$ positive offset on the same voltage measurement lead to a successful identification in $63.1\,\%$ of the cases. Similar results are found also when creating similar bad data ($2\,\%$ offset) on the other voltage measurements of the grid (successful identification in a range between $55\,\%$ and $65\,\%$ of the cases). Larger bad data with a $3\,\%$ offset are instead always correctly identified through the LNR test. These results thus further confirm that the proposed formulation keeps all the strengths of the conventional WLS formulation, while enabling the operation in theoretically unobservable scenarios.

G. Exemplary Operation during a 1-Day Time Window

As last verification of the performance of the proposed estimator, it may be interesting to see some exemplary results achievable when running the AF-WLS over the consecutive time steps of a day. In this case, the presented results are hence not derived from a MC analysis, but they are simply one-shot estimations obtained over different time steps when emulating a live operation of the estimator. To this purpose, typical daily profiles for residential and commercial loads, as well as for PV and wind generation, have been derived from the Atlantide project [35]. These profiles have a 15-minutes time resolution. To generate the variability of the power profiles within each cluster, it is assumed that the power of each customer can vary in the interval between $50\,\%$ and $150\,\%$ of the nominal allocation factor available at time t for residential loads and wind generation, whereas it varies between $75\,\%$ and $125\,\%$ for the commercial loads and PV generation (thus implicitly assuming that these categories of customers have a more uniform behavior within the cluster at a known time point). Tests have been run using the Case 2 measurement configuration.

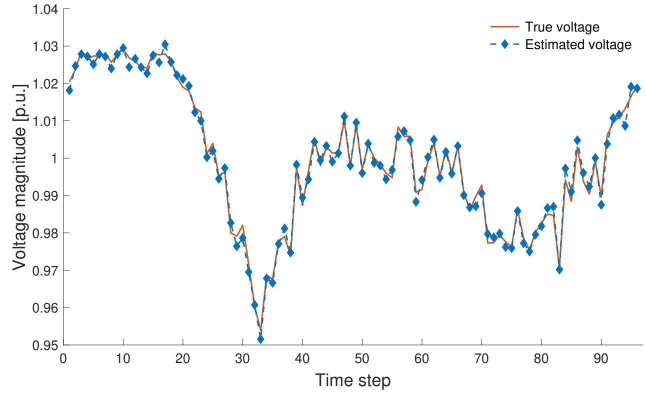

Fig. 10. Voltage magnitude estimation over the day for bus 111 (phase A).

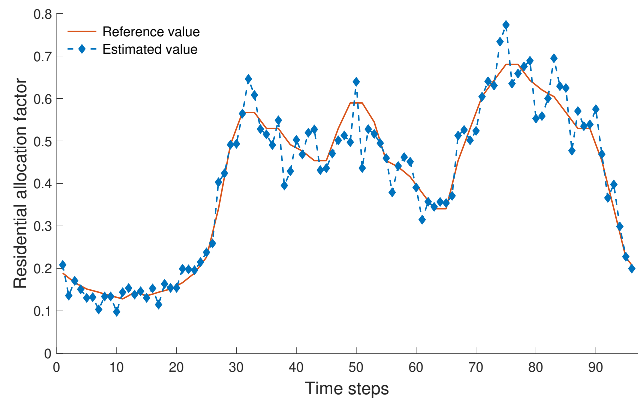

Fig. 11. Estimation of the allocation factor for residential loads over the day.

Figure 10 shows, as an example, the daily profile of the voltage magnitude estimation for bus 111 (phase A), which is a not measured bus. It is possible to observe that the proposed estimator allows following very closely the voltage variations over the day, with a maximum error that, in this specific example, is equal to $0.45\,\%$. Figure 11 shows instead, as an example, the reference scaling factors considered for the residential loads (with the characteristic peaks in the morning and in the evening) together with the estimated values of the corresponding allocation factor. In this case, it is possible to notice that estimation errors exist (coherently with what highlighted in the results of Fig. 5 and Fig. 6), but, again, the proposed estimator proves to be capable of providing reliable information also for the estimation of these parameters and to suitably track the changes occurring over time.

V. Conclusions

This paper presented a novel DSSE algorithm for unobservable distribution grids based on the classical WLS method. The proposed formulation relies on allocation factors to deal with the scarcity of measurement devices, it does not require the definition of any pseudo-measurements, it is relatively simple and it requires only few modifications to the conventional WLS SE model. Performed tests prove that the proposed estimator can work with very few measurements and can provide accurate and reliable SE results. Moreover, it allows avoiding some of the risks associated to the use of pseudo-measurements, above all in future scenarios where their definition may be troublesome. Finally, through the estimation of the allocation factors, the proposed estimator allows achieving a further level of detail in the knowledge of the operating conditions, which may be used, for example, for the monitoring of other unobservable areas or for other SE-related applications.

Appendix

In the following, the measurement functions used to express the field measurements with respect to the rectangular voltage state variables are given.

The voltage magnitude measurement $h_{V_{i,\phi}}$ at a generic bus i and phase ϕ is expressed as:

$$h_{V_{i,\phi}} = \sqrt{V_{ri,\phi}^2 + V_{xi,\phi}^2} \tag{16}$$

In case of PMUs, the measured voltage phasor is converted into rectangular coordinates and it is then easy to find:

$$h_{V_{ri,\phi}} = V_{ri,\phi} \qquad h_{V_{xi,\phi}} = V_{xi,\phi} \tag{17}$$

For the power and PMU current phasor measurements at the branch between the generic nodes i and k, both are converted into rectangular currents (see [28] for the conversion of power measurements into equivalent currents). Their measurement functions (for phase ϕ) are then expressed as:

$$h_{I_{ri,\phi}} = \sum_{\psi \in \Psi} \left[g_{ik,\phi\psi}(V_{ri,\psi} - V_{rk,\psi}) - b_{ik,\phi\psi}(V_{xi,\psi} - V_{xk,\psi}) \right] \tag{18}$$

$$h_{I_{xi,\phi}} = \sum_{\psi \in \Psi} \left[b_{ik,\phi\psi}(V_{ri,\psi} - V_{rk,\psi}) + g_{ik,\phi\psi}(V_{xi,\psi} - V_{xk,\psi}) \right] \tag{19}$$

where $\Psi = \{A, B, C\}$.

References

[1] M. Pau, M. Mirz, J. Dinkelbach, P. Mckeever, F. Ponci, and A. Monti, "A service oriented architecture for the digitalization and automation of distribution grids," *IEEE Access*, vol. 10, pp. 37 050–37 063, 2022.

[2] F. Ahmad, A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş, "Distribution system state estimation-a step towards smart grid," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 2659–2671, 2018.

[3] D. Della Giustina, M. Pau, P. A. Pegoraro, F. Ponci, and S. Sulis, "Electrical distribution system state estimation: measurement issues and challenges," *IEEE Instrumentation & Measurement Magazine*, vol. 17, no. 6, pp. 36–42, 2014.

[4] J. Liu, F. Ponci, A. Monti, C. Muscas, P. A. Pegoraro, and S. Sulis, "Optimal meter placement for robust measurement systems in active distribution grids," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 5, pp. 1096–1105, 2014.

[5] T. C. Xygkis and G. N. Korres, "Optimized measurement allocation for power distribution systems using mixed integer SDP," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 11, pp. 2967–2976, 2017.

[6] A. Mutanen, M. Ruska, S. Repo, and P. Jarventausta, "Customer classification and load profiling method for distribution systems," *IEEE Trans. Power Del.*, vol. 26, no. 3, pp. 1755–1763, 2011.

[7] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1888–1896, 2012.

[8] F. Adinolfi, F. D'Agostino, A. Morini, M. Saviozzi, and F. Silvestro, "Pseudo-measurements modeling using neural network and fourier decomposition for distribution state estimation," in *IEEE PES Innovative Smart Grid Technologies, Europe*, 2014, pp. 1–6.

[9] M. Ferdowsi, A. Benigni, A. Löwen, B. Zargar, A. Monti, and F. Ponci, "A scalable data-driven monitoring approach for distribution systems," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 5, pp. 1292–1305, 2015.

[10] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910–4920, 2019.

[11] B. Azimian, R. S. Biswas, S. Moshtagh, A. Pal, L. Tong, and G. Dasarathy, "State and topology estimation for unobservable distribution systems using deep neural networks," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–14, 2022.

[12] J.-H. Menke, N. Bornhorst, and M. Braun, "Distribution system monitoring for smart power grids with distributed generation using artificial neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 113, pp. 472–480, 2019.

[13] B. Habib, E. Isufi, W. v. Breda, A. Jongepier, and J. L. Cremer, "Deep statistical solver for distribution system state estimation," *IEEE Trans. Power Syst.*, pp. 1–12, 2023.

[14] A. S. Zamzam and N. D. Sidiropoulos, "Physics-aware neural networks for distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4347–4356, 2020.

[15] M.-Q. Tran, A. S. Zamzam, and P. H. Nguyen, "Enhancement of distribution system state estimation using pruned physics-aware neural networks," in *2021 IEEE Madrid PowerTech*, 2021, pp. 1–5.

[16] A. Akrami, S. Asif, and H. Mohsenian-Rad, "Sparse tracking state estimation for low-observable power distribution systems using d-pmus," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 551–564, 2022.

[17] G. Cavraro, J. Comden, E. Dall'Anese, and A. Bernstein, "Real-time distribution system state estimation with asynchronous measurements," *IEEE Transactions on Smart Grid*, vol. 13, no. 5, pp. 3813–3822, 2022.

[18] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2520–2530, 2020.

[19] S. Dahale and B. Natarajan, "Bayesian framework for multi-timescale state estimation in low-observable distribution systems," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4340–4351, 2022.

[20] I. D. Melo, M. O. Teixeira, and J. S. Mingorança, "Neutral-to-earth voltage (nev) and state estimation for unbalanced multiphase distribution systems based on an optimization model," *Electric Power Systems Research*, vol. 217, p. 109123, 2023.

[21] M. Pau, J. Dinkelbach, F. Ponci, and A. Monti, "A state estimation algorithm for the monitoring of distribution grids in absence of pseudo-measurements," in *NEIS 2020; Conference on Sustainable Energy Supply and Energy Storage Systems*, 2020, pp. 1–6.

[22] A. T. Sarić and A. Ranković, "Load reallocation based algorithm for state estimation in distribution networks with distributed generators," *Electric Power Systems Research*, vol. 84, no. 1, pp. 72–82, 2012.

[23] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*. CRC Press, 01 2004, vol. 24.

[24] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, "Uncertainty of voltage profile in PMU-based distribution system state estimation," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 5, pp. 988–998, 2016.

[25] M. Baran and A. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1601–1609, 1994.

[26] W.-M. Lin and J.-H. Teng, "Distribution fast decoupled state estimation by measurement pairing," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 143, pp. 43–48(5), January 1996.

[27] H. Wang and N. Schulz, "A revised branch current-based distribution system state estimation algorithm and meter placement impact," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 207–213, 2004.

[28] M. Pau, P. A. Pegoraro, and S. Sulis, "Efficient branch-current-based distribution system state estimation including synchronized measurements," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2419–2429, 2013.

[29] A. Angioni, T. Schlösser, F. Ponci, and A. Monti, "Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, 2016.
[30] W. H. Kersting, *Distribution System Modeling and Analysis*. BocaRaton, CRC Press, 2007.
[31] G. Valverde, A. T. Saric, and V. Terzija, "Stochastic monitoring of distribution networks including correlated input variables," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 246–255, 2013.
[32] J. Morales, L. Baringo, A. Conejo, and R. Minguez, "Probabilistic power flow with correlated wind sources," *Generation, Transmission & Distribution, IET*, vol. 4, no. 5, pp. 641–651, May 2010.
[33] JCGM, "Evaluation of data - guide to the expression of uncertainty in measurement," *JCGM 100:2008*, Sep. 2008.
[34] (2012) IEEE test feeder specifications. [Online]. Available: https://cmte.ieee.org/pes-testfeeders/resources/
[35] F. Pilo, G. Pisano, S. Scalari, D. Dal Canto, A. Testa, R. Langella, R. Caldon, and R. Turri, "ATLANTIDE — Digital archive of the Italian electric distribution reference networks," in *CIRED 2012 Workshop: Integration of Renewables into the Distribution Grid*, 2012, pp. 1–4.

Marco Pau (S'14-M'15) received the M.S. degree (cum laude) in electrical engineering and the Ph.D. degree in electronic engineering and computer science from the University of Cagliari, Cagliari, Italy, in 2011 and 2015, respectively. From 2015 to 2022, he was with the Institute for Automation of Complex Power Systems at RWTH Aachen University, Aachen, Germany, where he led the team for Distribution Grid Monitoring and Automation. Since 2022 he is a Senior Researcher at the Fraunhofer Institute for Energy Economics and Energy System Technology, where he leads the group for Decentralized Grid Automation. His research interests include the design of solutions for the monitoring and automation of power systems and for the smart management of distribution grids. Dr. Pau is an Associate Editor of the IEEE Transactions on Instrumentation and Measurement.

Paolo Attilio Pegoraro (Senior Member, IEEE) received the M.S. (*summa cum laude*) degree in telecommunication engineering and the Ph.D. degree in electronic and telecommunication engineering from the University of Padova, Padua, Italy, in 2001 and 2005, respectively.

From 2015 to 2018 he was an Assistant Professor with the Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy, where he is currently Associate Professor. He has authored or co-authored over 150 scientific papers. His current research interests include the development of new measurement techniques for power networks, with attention to synchronized measurements and state estimation.

Dr. Pegoraro is a member of IEEE IMS TC 39 (Measurements in Power Systems) and of IEC TC 38/WG 47. He is an Associate Editor of the IEEE Transactions on Instrumentation and Measurement and the General Chair of the IEEE International Workshop on Applied Measurements for Power Systems (AMPS).