Effects of Measurements and Pseudo-measurements Correlation in Distribution System State Estimation

Carlo Muscas, Member, IEEE, Marco Pau, Student Member, IEEE, Paolo Attilio Pegoraro, Member, IEEE, Sara Sulis, Member, IEEE.

Abstract—Distribution System State Estimation (DSSE) is one of the key elements of the monitoring activity of an active distribution network and is the basis for every control and management application. The DSSE relies on real measurements collected by the distributed measurement system and on other available information, mainly obtained from historical data, that help obtaining observability. This prior information is necessary to derive the so-called pseudo-measurements. Accurate input data are fundamental for an accurate estimation, as well as knowledge on possible correlation in the measured and pseudo-measured data. A degree of correlation can exist in the measured data, due to measurement devices, and among power consumptions or generations of some particular nodes. This paper presents an exhaustive analysis on the influence of correlations on the quality of the estimation. The importance of including correlation in the WLS estimation approach is discussed using both traditional and synchronized measurements. Results obtained on a 95-bus distribution network are presented and discussed.

Index Terms—Distribution systems, state estimation, distributed measurement systems, data correlation, measurement accuracy, Phasor Measurement Units.

I. INTRODUCTION

Distribution System State Estimation (DSSE) is becoming, in particular in the Smart Grid (SG) scenario, a fundamental tool of monitoring, control and management activities. In fact, control systems, such as Distribution Management Systems (DMSs), rely, for meaningful operation, on a knowledge of the state of the network, in terms of voltage profiles and current or power flows, as accurate as possible [1]. Distributed measurement systems are the underlying infrastructure of such monitoring activity, and DSSE methodologies represent the cornerstone to elaborate the measurements acquired from the field and translate them into the estimation of the operation point of the network. In this context, synchronized measurements and, in particular, the so-called synchrophasors given by Phasor Measurement Units (PMUs) are used along with voltage amplitude, current or active and reactive power measurements provided by traditional or new intelligent metering instruments. PMUs are becoming increasingly widespread in transmission systems and in state estimation [2]–[5], and they are expected to be widely used also in DSSE [6].

Since it is impractical and economically unfeasible to have a measurement device at each node of a distribution network, the DSSE reaches observability by relying on the a priori information on network behavior (the so-called pseudo-measurements, [7]), which is mainly represented by historical or forecast data on generators production and loads consumption. The better is the accuracy and completeness of the pseudo-measurements, the higher the quality of the final state estimation. Thus, it is fundamental to have a rich statistical description of loads and generators that can be included in the DSSE.

DSSE proposed in the literature are mainly based on weighted least squares (WLS) algorithms ([6]–[11]). In [7] the weights for the WLS were fixed and equal to 1 and 10 for real measurements and pseudo-measurements, respectively, while [6], [8]–[11] assume the inverse of the variances of measurements and pseudo-measurements as weights for measurement residuals.

However, the correlation among different measurements or pseudo-measurements is usually neglected and the analysis of its impact on the estimation results is lacking. There are few works in the literature, concerning both transmission and distribution systems, considering correlation. In [12], the correlation between loads has been considered in a probabilistic power flow, that uses the real measurements as constraints. In [13], the correlation among different measurements provided by a specific power meter device is used and it is shown that considering it into the WLS weighting matrix results in a more accurate estimation of the system state for transmission systems. Analogously, in [14], the possibility of including in the weighting matrix the correlation among different loads, generators and between active and reactive power injections has been considered. In [15], spatial correlation among loads or sources located in the same geographical area is used again to perform probabilistic power flows. In addition, as for transmission systems, in [16], a method to include possible correlations among transformer signals is presented.

In [17], the authors investigated the impact of the correlation among the pseudo-measurements at some nodes of a distribution network on the accuracy of a branch current DSSE algorithm. The importance of refining the load models to be used in the DSSE and matching the statistical behavior of the pseudo-measurements was discussed, using both traditional and synchronized measurements.

In this paper, further steps in the investigation of the impact of such correlation onto the DSSE accuracy are performed. A deeper analysis is carried out to study the impact of the correlation inside different distribution subnetworks. The importance of modelling the type and characteristics of the
loads is proved. Furthermore, the correlation among the real measurements is discussed, both for traditional measurement devices and for PMUs. Methodologies to include such correlation in the DSSEs are reported along with the description of the correlation sources inside a measurement station. The impact on the network state estimates is assessed considering a 95-bus test network.

II. DISTRIBUTION SYSTEM STATE ESTIMATION

Distribution networks have a few measurement points and, as a consequence, pseudo-measurements are necessary to know the state of the system. Different state variables can be considered for the estimation algorithm, since different quantities (voltages, currents, powers) and forms (polar or rectangular) can be used, depending on the needs.

In this paper, the branch-current based estimator proposed in [6] is adopted. This algorithm will be referred here as BC-DSSE. It is worth noting that the considerations derived from the tests on BC-DSSE are expected to hold even for voltage state DSSE algorithms based on WLS (like [7]), since, as demonstrated in [18], the accuracy of the two classes is similar. The BC-DSSE has been chosen because it is the most recent DSSE based on WLS proposed in the literature, it is really fast and is designed to allow the inclusion of the synchronized phasor measurements.

The DSSE relies on a general measurement model that can be represented as:

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

where \( \mathbf{z} = [z_1 \ldots z_M]^T \) is the vector of the measurements gathered from the network and of the chosen pseudo-measurements (\( M \) is the total number of available measurements); \( \mathbf{h} = [h_1 \ldots h_M]^T \) is, in general, the vector of non-linear measurement functions (depending on the type of measurements); \( \mathbf{x} = [x_1 \ldots x_N]^T \) is the state vector, composed of \( N \) state variables. The measurement noise vector \( \mathbf{e} \) is usually assumed to be composed by independent zero mean Gaussian variables. The state vector of the chosen DSSE, as aforementioned, includes the branch currents of all the \( N_{br} \) network branches and the voltage at a reference bus, for instance the slack, in rectangular coordinates. In particular the \( \mathbf{x} \) becomes equal to 

$$\begin{bmatrix} \mathbf{i}_{\text{slack}}^T \mathbf{i}_{\text{Nbr}}^T \mathbf{i}_{\text{Nbr}}^T \mathbf{i}_{\text{Nbr}}^T \mathbf{i}_{\text{Nbr}}^T \end{bmatrix}^T$$

when synchronized phasor measurements are present and to 

$$\begin{bmatrix} \mathbf{v}_{\text{slack}}^T \mathbf{i}_{\text{Nbr}}^T \mathbf{i}_{\text{Nbr}}^T \mathbf{i}_{\text{Nbr}}^T \mathbf{i}_{\text{Nbr}}^T \end{bmatrix}^T$$

when there are only traditional measurements. In fact, when no synchronized measurements are present, the slack bus voltage phase angle has to be assumed equal to zero since it serves as reference for the other buses phase angles. When synchronized measurements are present, it is instead possible to measure the absolute phase angles with respect to the coordinated universal time (UTC), given by global positioning system (GPS), or other synchronization sources (see for instance [19], [20]). The pseudo-measurements usually consist in a priori information about the consumption of the loads or the production of the generators. The real measurements can be voltage amplitude, branch current and current injection amplitude, power flow or power injection, voltage and current phasor measurements, depending on the installed devices. In general the distributed measurement system involves heterogeneous measurement types and different measurement accuracies.

Even if in the literature (see for example [6], [11]) the covariance matrix is usually considered diagonal, the general expression of the covariance that includes correlation can be expressed as:

$$\Sigma_{\mathbf{x}} = \begin{bmatrix} \sigma_{z_1}^2 & \ldots & \rho_{1,j} \sigma_{z_1} \sigma_{z_j} & \ldots & \rho_{1,M} \sigma_{z_1} \sigma_{z_M} \\ \vdots & & \ddots & & \vdots \\ \rho_{i,j} \sigma_{z_i} \sigma_{z_j} & \ldots & \ddots & \ldots & \vdots \\ \vdots & & \ddots & & \ddots \\ \rho_{i,M} \sigma_{z_i} \sigma_{z_M} & \ldots & \ldots & \ldots & \sigma_{z_M}^2 \end{bmatrix}$$  \hspace{1cm} (2)

where \( \sigma_{z_j} \) is the standard deviation of the \( j \)-th measurement and \( \rho_{i,j} \) is the correlation factor between measurements \( i \) and \( j \). Measurements have a lower standard deviation than pseudo-measurements and this reflects the higher accuracy of real-time measurements over historical data.

In the WLS approach the estimation of the state \( \mathbf{x} \) is usually obtained by the minimization of the weighted sum of the squares of the residuals:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \sum_{i=1}^{M} w_i (z_i - h_i(\mathbf{x}))^2$$  \hspace{1cm} (3)

$$= (\mathbf{z} - \mathbf{h}(\mathbf{x}))^T \mathbf{W} (\mathbf{z} - \mathbf{h}(\mathbf{x}))$$

where \( w_i \) is the weight associated to the \( i \)-th measurement. In the second term of (3) the general matrix form of the WLS is reported, where \( \mathbf{W} = \Sigma_{\mathbf{e}}^{-1} \) is the weighting matrix.

Because of the non linearity of the measurement functions, for the solution of (3), a Gauss-Newton iterative approach is used. At each iteration \( n \), \( \Delta \mathbf{x} = \mathbf{x}_n - \mathbf{x}_{n-1} \) is computed starting from the measurement residuals, by solving the so-called normal equations:

$$\mathbf{H}^T \mathbf{W} \mathbf{H} \Delta \mathbf{x} = \mathbf{H}^T \mathbf{W} (\mathbf{z} - \mathbf{h}(\mathbf{x}_{n-1}))$$  \hspace{1cm} (4)

where \( \mathbf{H} \) is the Jacobian of the measurement functions and is, in general, function of the state \( \mathbf{x}_{n-1} \) of the previous iteration. \( \mathbf{G} = \mathbf{H}^T \mathbf{W} \mathbf{H} \) is the gain matrix that has to be inverted to find the solution of the system.

In the BC-DSSE approach, at each iteration, before solving (4), all the power measurements and pseudo-measurements are translated into equivalent current measurements to simplify the measurement model for power measurements (see [9] and [11] for details) thus allowing a straightforward integration in the estimator. Then, the state variation \( \Delta \mathbf{x} \) is computed and all the voltages are calculated starting from the slack bus voltage estimate and using the voltage drops associated with the branch currents (forward sweep step). Such computation can be synthetically expressed in a matrix multiplication (for each iteration \( n \) ) as:

$$\mathbf{v}_n = \mathbf{Z}_{\text{paths}} \mathbf{x}_n$$  \hspace{1cm} (5)

where \( \mathbf{v} = [v_1 \ldots v_N]^T \) is the complex voltages vector and \( \mathbf{Z}_{\text{paths}} \) is the matrix that contains, for each row \( i \), the path of the branch impedances that links \( v_i \) to \( v_{\text{slack}} \). The calculated voltages are also used to compute the measurement vector
for the next iteration and the procedure is repeated until the desired convergence is reached.

Looking at (4), it is clear that the presence of the correlation reduces the sparsity of the weighting matrix $W$ and, as a consequence, the speed in the inversion of $G$. It is then important to understand when it is relevant to include correlation to enhance the estimation performance and when, on the other end, the estimation can be kept accurate without degrading the speed of the estimation process.

III. CORRELATION OF THE MEASUREMENTS

Measurement values given by different devices can be reasonably considered independent. As a consequence, correlation is assumed for voltage amplitudes, current amplitudes, voltage and current phasors, real and reactive power flows or power injection measurements collected by different instruments. On the contrary, correlation among the different measurements given by the same device has to be taken into account, for both traditional measurement units and PMUs. Besides, as pointed out in [17], correlation can arise among different loads or generators and between active and reactive powers.

A. Correlation in traditional measurement devices

A traditional measurement device is assumed, for the following analysis and test results in the Section IV, to be able to measure both voltage amplitude and power flows (in particular, two measurement channels are considered for active and reactive power flows). The correlation among measurements provided by such devices can arise, specifically, among voltage magnitude ($V$) and active and reactive power measurement ($P$ and $Q$).

As described in [13], digital power meter devices usually provide voltage and power measurements starting from the direct measurements of voltage and current magnitude ($V$ and $I$) and voltage-current phase angle ($\phi = \theta_v - \theta_i$) and computing the active and reactive power indirectly. Correlation is thus introduced in the elaboration and it is possible to take into account the independent measurements uncertainties to achieve variances and covariances of the different measurements provided by the device.

If, for simplicity, only the fundamental components are taken into account, the measured $P$ and $Q$ can be expressed as functions of $V$, $I$ and $\phi$ as:

$$P = P(V, I, \phi) = VI \cos(\phi)$$
$$Q = Q(V, I, \phi) = VI \sin(\phi)$$

and grouped in a vector function relationship:

$$f(V, I, \phi) = \begin{pmatrix} V \\ P(V, I, \phi) \\ Q(V, I, \phi) \end{pmatrix}$$

(7)

Indicating the standard deviations of voltage and current magnitudes and phase-angle as $\sigma_V$, $\sigma_I$, $\sigma_\phi$, respectively, the resulting covariance matrix can be calculated by means of matrix multiplication (where $z_{v\phi} = (V \ P \ Q)^T$ is the corresponding measurement deviations sub-vector) as:

$$E [\Delta z_{v\phi} \Delta z_{v\phi}^T] = \begin{bmatrix} \sigma_V^2 & \sigma_{VP} & \sigma_{VQ} \\ \sigma_{PV} & \sigma_P^2 & \sigma_{PQ} \\ \sigma_{QV} & \sigma_{QP} & \sigma_Q^2 \end{bmatrix}$$

(8)

where $\sigma_{v\phi}$ represents the Jacobian of the function vector (7) with respect to the independent measurements.

By substituting the derivatives and solving the matrix multiplication (see the Appendix, for details), the following variances can be found:

$$\sigma_P^2 = \sigma_V^2 I^2 \cos^2 \phi + \sigma_I^2 V^2 I^2 \sin^2 \phi + \sigma_\phi^2 V^2 \cos^2 \phi$$

(9)

$$\sigma_Q^2 = \sigma_V^2 I^2 \sin^2 \phi + \sigma_I^2 V^2 I^2 \cos^2 \phi + \sigma_\phi^2 V^2 \sin^2 \phi$$

(10)

As for the covariances terms, the following results can be obtained:

$$\sigma_{VP} = \sigma_V^2 I \cos \phi$$

(11)

$$\sigma_{VQ} = \sigma_V^2 I \sin \phi$$

(12)

$$\sigma_{PQ} = \frac{1}{2} (\sigma_V^2 I^2 \sin 2\phi - \sigma_\phi^2 V^2 I^2 \sin 2\phi + \sigma_\phi^2 V^2 \sin 2\phi)$$

(13)

By considering that information provided by the power meter only refer to voltage and power measurements, it is necessary to express the equations in (9)-(13) in terms of the above mentioned measurements and of the independent measurement uncertainties. Since the uncertainty is usually expressed in terms of relative values, it is simpler to present the variances terms in (9-10) divided by the squared measured quantities, thus obtaining:

$$\frac{\sigma_P^2}{P^2} = \frac{\sigma_V^2}{V^2} + \frac{\sigma_I^2}{I^2} + \sigma_\phi^2 \tan^2 \phi$$

(14)

$$\frac{\sigma_Q^2}{Q^2} = \frac{\sigma_V^2}{V^2} + \frac{\sigma_I^2}{I^2} + \frac{\sigma_\phi^2}{V^2} \cot^2 \phi$$

(15)

Instead, as for the covariances terms, equations in (11)-(13) lead to:

$$\frac{\sigma_{VP}}{VP} = \frac{\sigma_V^2}{V^2}$$

(16)

$$\frac{\sigma_{VQ}}{VQ} = \frac{\sigma_V^2}{V^2}$$

(17)

$$\frac{\sigma_{PQ}}{PQ} = \frac{\sigma_V^2}{V^2} + \frac{\sigma_I^2}{I^2} - \sigma_\phi^2$$

(18)

If equivalent current measurements are used in place of the power flow measurements, as in the BC-DSSE, an equivalent covariance matrix must be calculated to include covariances in the DSSE. The relationship describing the equivalent current measurements (see [9] for further details) is:

$$i^r + ji^s = \frac{(P - jQ)(v^r + jv^s)}{V^2}$$

(19)
If voltage magnitude $V \approx 1$ p.u., the following equations hold for the equivalent current measurements:

\[ i^r \approx P v^r + Q v^x \]  \hspace{1cm} (20)

\[ i^x \approx P v^x - Q v^r \]  \hspace{1cm} (21)

In a single phase framework, if only traditional measurements are available on the network, the slack bus phase angle is chosen as reference for all the other phase angles and is usually set equal to zero. In this case, considering the small voltage drops along the lines of a distribution network, the voltage phase angles are really close to zero, thus we have: $v^r \approx V$ and $v^x \approx 0$. With these assumptions, (20) and (21) are $i^r \approx P$ and $i^x \approx -Q$, then the covariance matrix including the equivalent current measurements, will become:

\[
\begin{pmatrix}
\sigma_i^2_V & \sigma_{iV} & \sigma_{iV} \\
\sigma_{iV} & \sigma_{iV}^2 & \sigma_{iV}^2 \\
\sigma_{iV} & \sigma_{iV} & \sigma_{iV}^2
\end{pmatrix}
\approx
\begin{pmatrix}
\sigma_P^2 & \sigma_{PV} & -\sigma_{VQ} \\
\sigma_{PV} & \sigma_P^2 & -\sigma_{QP} \\
-\sigma_{QV} & -\sigma_{QP} & \sigma_Q^2
\end{pmatrix}
\]  \hspace{1cm} (22)

In this case only a change of sign in the covariances including $i^x$ is required. Instead, if PMU measurements are also available or if a three-phase framework is considered, the assumption of voltage phase angles close to zero does not hold any more. Thus, the covariance matrix taking into account the equivalent measurements should be evaluated using the formulation of the error propagation (similarly to the one used in (8)) and referring to equations (20) and (21).

It is worth noting that the aforementioned approximations are used only for the weighting matrix computation and, thus, for a faster iteration of the algorithm. Tests have been performed and confirm that, as illustrated by several papers (see for example [8], [18]), the comparison between results obtained with a classic polar voltage estimator and a formulation based on branch currents and equivalent current measurements shows that approximations introduced in the equivalent measurements do not significantly affect the estimation accuracy.

In Section IV, the described covariances will be specified for possible applications, to understand typical correlation levels and to show their impact on the estimations.

B. Correlation in PMU measurements

PMU measurements include voltage phasor at the node and current phasors of the branches adjacent to the node where PMU is installed. The number of channels depends on the PMU and in the tests it has been assumed that two currents are measured, thus allowing comparison between PMU and traditional measurement station. Generally speaking, correlation can arise between different measurement channels and between amplitude and phase measurement in the same channel. The measurement of voltage and current amplitudes in the same PMU can be usually considered decorrelated. In fact, starting from device specifications and given normal operative conditions, no significant common uncertainty causes can be identified. On the other side, phase angle measurements are based on the same synchronization. For this reason, in the following, the effect of a common reference time in the PMU is considered and a method to include this correlation in the estimator is presented.

To evaluate the uncertainty of the phase-angle measurements performed by a PMU, the sum of three contributions can be considered:

\[ \epsilon_A = \epsilon_{TA} + \epsilon_{SA} + \epsilon_{th} \]  \hspace{1cm} (23)

\[ \epsilon_B = \epsilon_{TB} + \epsilon_{SB} + \epsilon_{th} \]  \hspace{1cm} (24)

where $\epsilon_A$ and $\epsilon_B$ are the phase deviations on two different channels A and B, $\epsilon_{TA}$, $\epsilon_{TB}$ (i = A, B) is the effect of the transducer phase error (or of the residual uncertainty, if a compensation of the transducer errors is performed), $\epsilon_{SA}$, $\epsilon_{SB}$ (i = A, B) is the phase deviation introduced by the inner PMU signal conditioning system, and $\epsilon_{th}$ is the phase deviation related to timebase deviation, that is to the achieved synchronization level. All the above variables can be considered zero-mean. Calculating the covariance of $\epsilon_A$ and $\epsilon_B$:

\[
E[\epsilon_A \epsilon_B] = E[(\epsilon_{TA} + \epsilon_{SA} + \epsilon_{th})(\epsilon_{TB} + \epsilon_{SB} + \epsilon_{th})]
\]

\[
\approx E[\epsilon_A^2 \epsilon_B^2] = \sigma_{th}^2
\]  \hspace{1cm} (25)

where the last equality is obtained by considering decorrelation among $\epsilon_{TA}$, $\epsilon_{TB}$, $\epsilon_{SA}$, and $\epsilon_{SB}$, and between $\epsilon_{th}$ and the other components. To obtain the correlation factor, the standard deviations should be considered:

\[
\sigma_i^2 \triangleq E[\epsilon_i^2] = \sigma_{th}^2 + \sigma_{TA}^2 + \sigma_{SB}^2
\]  \hspace{1cm} (26)

where $i = A, B$ and the second term is obtained by considering the additive terms as decorrelated, as aforementioned. If the two channels can be assumed equal, as it appears reasonable, $\sigma_{SA} = \sigma_{SB}$. Finally, the correlation factor can be computed merging (25) and (26) as:

\[
\rho_{A,B} = \frac{E[\epsilon_A \epsilon_B]}{\sigma_A \sigma_B} = \frac{\sigma_{th}^2}{\sqrt{\sigma_{TA}^2 + \sigma_{SA}^2 + \sigma_{th}^2} \sqrt{\sigma_{TB}^2 + \sigma_{SB}^2 + \sigma_{th}^2}}
\]

\[
= \frac{1}{\sqrt{1 + (\sigma_{TA}^2 + \sigma_{SA}^2)/\sigma_{th}^2} \sqrt{1 + (\sigma_{TB}^2 + \sigma_{SB}^2)/\sigma_{th}^2}}
\]  \hspace{1cm} (27)

Such equation, as is intuitive, shows how the correlation factor strictly depends on the ratio between the independent contributions to the phase error and the common ones, and can be further simplified if $\sigma_{TA} = \sigma_{TB}$ (for example, when analogous transformers and similar currents are considered).

Using the aforementioned assumptions and knowing the device specifications for each contribution, usually reported by the PMU datasheet (see for instance [21]), standard deviations can be computed. It is worth recalling that standard deviations can, obviously, depend on the measured values.

In Section IV, correlation factors will be calculated referring on realistic applications and used to test the impact of PMU correlation on the DSSE accuracy.

C. Correlation in pseudo-measurement information

The number of pseudo-measurements in DSSE is high with respect to real measurements typically available in a distribution system. A given degree of correlation can be present, in sub-areas of the distribution network, among power consumption at different nodes of the network (inter-node
correlation) due to several causes. In fact, loads can present similarity, in particular on a geographical basis. Weather conditions, energy cost or network management operation can influence the inter-node correlation.

Besides, correlation can become relevant for the power injected by distributed generation (DG) or renewable energy sources. In fact, weather conditions or management interventions (such as reactive generation control or active power curtailment) can have a strong effect in correlating generators or prosumers that belong to the same geographical area.

In a DSSE perspective, it is important to have a measure of the correlation factor between two pseudo-measurements. Such factor is, from the point of view of a priori modelling, a relative concept: the correlation level can change depending on available information and accuracy of unmonitored loads description. For instance, if the knowledge about load absorption is very accurate, even if the behaviour of two loads is very similar, the correlation to be inserted in the model and, thus, in the DSSE can be low, because it describes only the residual random fluctuation around the known values. In DSSE, the temporal load profiles are usually used and, in this context, correlation should describe the relationship among the variations of the loads with respect to the expected values.

Correlation can be assumed also between active and reactive power consumption at the same bus (intra-node correlation) considering that a high degree of knowledge of the power factor is a common assumption for the MV loads. These types of correlations have been considered in this paper for the analysis aimed at understanding their impact on DSSE based on different measurement systems.

It is important to recall that when different levels of correlation are present among different quantities, all the indirect cross-correlation factors have to be considered, when building the covariance matrix of the measurements. As an example, if the power consumptions of two loads are correlated and active and reactive powers of each node are also related, spurious correlation arises also between the active power of one node and the reactive one of the other. In the following tests, spurious correlation factors are computed by multiplication of the different contributing terms.

IV. TESTS AND RESULTS

In order to analyse the impact of the different correlated input data onto the accuracy of the DSSE, several tests have been performed on a 95-bus network (Fig.1). This is a symmetrical radial system, with two sources of DG\(^1\). Nominal data of the network were obtained from [22]. Simulations have been carried out by means of Matlab.

To assess the uncertainty in the DSSE, several Monte Carlo trials have been used. In each trial, the absorbed and generated powers have been randomly chosen from the respective distributions, the power flow has been computed to obtain the reference values, and the random measurement errors have been extracted considering the measurement uncertainty and taking into account the assumed correlations. In particular, for the presented tests, the following assumptions are considered:

- Number of Monte Carlo trials \(N_{MC} = 50000\);
- A maximum deviation of 50\% (Gaussian distribution) with respect to the pseudo-measured values for the active and reactive power injections on all the load and generation nodes;
- Normal distribution for the real measurements with a standard deviation equal to a third of its accuracy value.

The normal probability distribution, along with a coverage factor equal to 3, has been assumed for the tests, since it is a common assumption for this kind of data (see, for example, [22], [23]).

Simulations have been carried out considering a possible configuration of the measurement system. In particular, measurement devices have been located in different points of common coupling of the network. Each measurement point has one voltage measurement on the node and two flow measurements (current or power) on adjacent branches. Three possible measurement configurations have been considered:

- Type A: voltage magnitude measurement on the node and current magnitude measurements on the branches;
- Type B: voltage magnitude measurement on the node and active and reactive power measurements on the branches;
- Type C: phasor measurements provided by PMUs for both the voltage of the node and the current of the branches.

Table I reports the positions of the chosen measurement points in the considered measurement system. Such choice is only based on technically sound reasoning, without using optimal meter placement techniques.

---

\(^{1}\)As for the numeration of the branches, each branch index is given by the node number of its end node (the largest one), decreased by one.

<table>
<thead>
<tr>
<th>Nodes</th>
<th>1</th>
<th>11</th>
<th>37</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branches</td>
<td>1 (1-2)</td>
<td>3 (1-4)</td>
<td>11 (11-12)</td>
<td>28 (11-29)</td>
</tr>
<tr>
<td>(Start-End node)</td>
<td>36 (37-36)</td>
<td>40 (37-41)</td>
<td>60 (60-61)</td>
<td>76 (60-77)</td>
</tr>
</tbody>
</table>

In addition to the selected real measurements, the pseudo-measurements in all the load and generation nodes are considered to be known. These, together with the zero-injection nodes (here included as virtual measurements with very large weight) allow to obtain the observability of the network. It is
worth noting that such a low number of real measurements is quite realistic in distribution networks (see for example [1], [7], [11], [22]) where, often, measurement devices could be available only in the primary substations.

In the presented tests, the impact of including the correlation into the estimator model is analysed assessing the Root Mean Square Errors (RMSE) of the estimated quantities over the $N_{MC}$ trials, according to the following:

$$RMSE_y = \sqrt{\frac{\sum_{i=1}^{N_{MC}} (y_i - \hat{y}_i)^2}{N_{MC}}}$$

(28)

where $y_i$ and $\hat{y}_i$ are the true and estimated value of the considered quantity, respectively.

### A. Traditional measurement correlation tests

As shown in Section III-A, the correlation among the measurements provided by a power meter strictly depends on the measured quantities themselves. Thus, it is not possible to simply define a specific value of the correlation factor to be used in the simulations. Tests, in this context, are hence performed considering the real behaviour of the power meter. First of all, measurements of voltage and current are extracted, starting from the reference values: in the tests the accuracy values are $1\%$ for the magnitude measurements and $5.8 \times 10^{-3}$ rad for the phase-angle measurements (as indicated in [24], [25] for the 0.5-class transformers and assuming phase-angle errors totally caused by transducers). From these data, the corresponding measurements of voltage and power provided by the power meter are calculated.

Performed tests have shown the presence of correlation factors varying from 0.3 to 0.8 depending on the involved quantities and the considered measurement point. In general, results do not highlight large improvements on the estimation accuracy when correlations are duly considered. Table II reports a summary of the outcomes, the average of the RMSE among all the branches or nodes are showed. It is important to highlight that, in particular for branch currents, the effect of correlation is local and thus specific results can be slightly more evident than in Table II: considering for instance current magnitude of branch 7, the percent RMSE can pass from more evident than in Table II: considering for instance current magnitude of branch 7, the percent RMSE can pass from $0.018^\circ = 0.31 \cdot 10^{-3}$ rad; a maximum deviation equal to 0.03 $\%$ for the magnitude measurements.

Using the assumed values for the PMU phase deviations, if 0.5-class transducers (with maximum deviation of $5.8 \times 10^{-3}$ rad as indicated in [24], [25]) are considered for both channels, a correlation factor $\rho_{A,B} = 2.8 \times 10^{-3}$, thus negligible, is obtained. Instead, if the transducer error can be neglected, due to fine compensation, the correlation factor becomes $\rho_{A,B} \simeq 0.26$.

Results obtained using this latter correlation factor do not show any improvement on the estimation accuracy: almost the same results have been obtained both neglecting and including correlations into the DSSE. A maximum deviation equal to 0.03 $\%$ for the magnitude measurements.

Results obtained using this latter correlation factor do not show any improvement on the estimation accuracy: almost the same results have been obtained both neglecting and including correlations into the DSSE. A maximum deviation equal to 0.03 $\%$ for the magnitude measurements.

### B. PMU measurement correlation tests

Tests on PMU measurement correlation have been performed taking into account the considerations reported in Section III-B. A first analysis has been performed using the data sheet of a commercial PMU [21]. The following assumptions on the accuracy values have been made:

- A maximum deviation of the system phase error $\epsilon_S$ equal to $0.03^\circ = 0.52 \cdot 10^{-3}$ rad, and an accuracy of the term associated to the time base error equal to $1 \mu s$.

In order to stress this aspect, also following aforementioned considerations, another test has been performed. In this case, a lower accuracy (0.7 $\%$ for magnitude measurements and 0.7 $\%$ for phase-angle measurements) has been chosen and a very high correlation degree (0.9) has been assumed. Table III shows global results for the network by means of the average of the RMSEs among all the branches or nodes. Also in this case, even though an extremely high correlation level is considered, including correlation into the estimator does not imply large improvements. In fact, even in such an extreme correlation scenario, the enhancements are not relevant for voltage estimations, while for currents they are analogous to
those of previous Section. Therefore, in realistic conditions, neglecting PMU correlations in the estimator model does not seem to lead to a significant degrade of the estimation quality.

C. Pseudo-measurement correlation tests

As for the pseudo-measurements, as mentioned in Section III-C, two different kinds of correlation can be analysed. The first one is an inter-node correlation, that is the correlation between the powers of two different nodes of the network. The second one is an intra-node correlation, that is the correlation between the active and the reactive powers into the same node.

Accuracy values of the measurement devices used in the following tests are: 1% for the voltage and current magnitudes, 3% for the active and reactive powers (Type A and B configurations), and 0.7% and 0.7 crad for magnitude and phase-angle measurements of PMUs (Type C), respectively.

As for the inter-node correlations, following the assumptions reported in [14], different groups of correlated loads have been identified:

- zone A: loads included between nodes 12 and 27;
- zone B: loads included between nodes 38 and 54;
- zone C: loads included between nodes 62 and 76;
- zone D: loads included between nodes 81 and 94.

For each group, a correlation among the active powers of the load nodes, with correlation factor equal to 0.8, is considered. In addition, a strong correlation (correlation factor 0.9) is assumed between DG on nodes 28 and 95.

A first series of tests has been performed taking into account correlation in only one of the above mentioned zones and between the generation nodes.

Figures 2 and 3 show the results for the branch current estimations in presence of inter-node correlation among the loads of zone A and in the case of Type B measurement devices. It is possible to see that a relevant improvement in the accuracy of the current estimation can be found when the correlations are included into the weighting matrix. The benefits are localized on the branches near to the correlated nodes. As for the voltage estimations, better estimations can be found, but with lower enhancement than for the currents. Similar results have been obtained also for the other measurement devices (Type A and Type C) and considering the correlation in the other zones.

Figure 4 shows the results obtained for the current magnitude estimations assuming correlation among the loads for all the zones and between the generators. In this case, as expected, results are similar to the previous case, but with a larger number of branches involved in estimation enhancement.

As for the intra-node correlation, correlation between active and reactive power of the same node has been studied. Several correlation factors have been tested. Results show that, in this case, the impact is mainly on the phase-angle estimations and that it is strongly dependent on the type of measurement devices available on the network. In fact, the intra-node correlation provides additional information about the power factor of the power injections. Thus, it is particularly useful to integrate data coming from measurement devices which do not provide any information about the phase-angles of the electrical quantities (or provide it with poor accuracy). Figures 5 and 6 show the results for the current phase-angle estimations in case of Type A and Type C measurements, respectively, when intra-node correlation, with correlation factor equal to 0.8 in all the load nodes, is considered. In particular, these results refer to power factors of the nodes supposed basically known and with small variations. Such assumption is very common in this kind of study and refers to usual operating conditions of the networks.

It is possible to observe that, when voltage and current magnitude measurements are used, an improvement in the current phase-angle estimations of the branches near to the
measurement points can be found. On the contrary, in case of PMU measurements, since the devices already provide accurate information about the phase-angles, there are no significant differences in the estimation results.

Tests including both inter-node and intra-node correlation have also been carried out. In this case, besides the assumed correlations, also spurious correlations arise (for example between the active power of a load and the reactive power of a correlated node when both are characterized also by intra-node correlation). Results show a cumulative effect of the enhancements associated to the different types of correlation. In Fig. 7, results obtained with inter-node correlation in all the zones and intra-node correlation in all the loads are reported. From the comparison between Figs. 5 and 7, it is possible to observe the additional contribution given by the inter-node correlation. Results show a cumulative effect of the enhancements associated to the different types of correlation.

In Fig. 7, results obtained with inter-node correlation in all the zones and intra-node correlation in all the loads are reported. From the comparison between Figs. 5 and 7, it is possible to observe the additional contribution given by the inter-node correlation. It is also important to highlight that, in such a situation, because of the large number of correlated data, more significant improvements are also obtained for the voltage estimations. In Fig. 8 results obtained for the voltage magnitude estimations are reported: improvements exist above all in the nodes farther from measurement points.

Results obtained until now have been achieved considering pseudo-measurements for all the load or generation nodes of the network. In order to analyse the impact of correlation also when some pseudo-measurements are replaced by real measurements, another test has been performed supposing to have real power measurements in the generation nodes and in the load nodes 3, 13, 53, 66 and 81. Thus, in such nodes the correlation has been removed while in the pseudo-measured nodes both intra-node and inter-node correlation has been considered. Fig. 9 clearly shows that the impact of correlation is still evident in all the branches near to the correlated nodes. It is possible to observe that similar results have been obtained, obviously with the exception of the branches adjacent to the nodes monitored with real measurements, where a better estimation is now achievable. In the following, the reduced measurement systems is again assumed for the tests.

The analysis on the impact of correlated pseudo-measurements, has also been extended to study the effects obtained with different levels of correlation. Tests have been performed taking into account inter-node correlation for all the load groups, correlation between the generators and intra-node correlation. Table IV shows results for the current magnitude estimations (Type B measurements) on branches 25, 39 and 75 when the direct correlations are reduced to 25%, 50% and 75% of the reference case. Obviously, the impact is larger when higher correlation factors are used. Finally, it is worth nothing that the impact is non-linear, also due to the presence of spurious correlations.

Some tests have been performed also to understand which is the effect of a possible underestimation or overestimation of the correlation factors, in order to simulate situations in which correlation degrees are not perfectly known. The aim of this study is to analyse what can happen in a realistic environment.
In this paper, the impact of the input data correlation on DSSE results has been presented and analysed. Different types of measurement devices and different kind of possible correlations have been considered. Correlations in traditional measurement devices, in PMU measurements, and in pseudo-measured data have been analysed and discussed. Different results have been obtained depending on the kind of correlation. As for the correlation among real measurements, outcomes show that only slight improvements can be obtained. In case of correlated pseudo-measurements, instead, results highlight that a proper modeling of the measurement covariances, obtained by including possible correlations in the weighting matrix of WLS, is really useful to improve the estimation accuracy. Moreover, tests show that important benefits can also be obtained when correlation factors are not perfectly known or estimated. This is an important feature for realistic situations where the knowledge (or estimation) of the possible correlations could be affected by a certain degree of uncertainty.

**APPENDIX**

The vector function \( f \) (defined in (7)) relates the independent measurements vector \( z_{\psi \phi} = (V \phi I)^T \) to the dependent measurements one \( z_{\text{pq}} = (V P Q)^T \). Its Jacobian matrix \( \frac{\partial f}{\partial z_{\psi \phi}} \) can be expanded as follows:

\[
\frac{\partial f}{\partial z_{\psi \phi}} = \begin{bmatrix}
\frac{\partial V}{\partial V} & \frac{\partial V}{\partial V} & \frac{\partial V}{\partial V} \\
\frac{\partial \phi}{\partial \phi} & \frac{\partial \phi}{\partial \phi} & \frac{\partial \phi}{\partial \phi} \\
\frac{\partial I}{\partial I} & \frac{\partial I}{\partial I} & \frac{\partial I}{\partial I} \\
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & V \sin \phi & V \cos \phi \\
I \sin \phi & V I \cos \phi & V \sin \phi \\
\end{bmatrix}
\]

and thus (9)-(13) are obtained by the following uncertainty propagation:

\[
E \left[ \Delta z_{\text{pq}} \Delta z_{\text{pq}}^T \right] =
\begin{bmatrix}
\sigma_v^2 & \sigma_v \sigma_p & \sigma_v \sigma_q \\
\sigma_p \sigma_v & \sigma_p^2 & \sigma_p \sigma_q \\
\sigma_q \sigma_v & \sigma_q \sigma_p & \sigma_q^2 \\
\end{bmatrix} \times
\begin{bmatrix}
1 & 0 & 0 \\
0 & V \sin \phi & V \cos \phi \\
I \sin \phi & V I \cos \phi & V \sin \phi \\
\end{bmatrix} \times
\begin{bmatrix}
\sigma_v^2 & \sigma_v \sigma_p & \sigma_v \sigma_q \\
\sigma_p \sigma_v & \sigma_p^2 & \sigma_p \sigma_q \\
\sigma_q \sigma_v & \sigma_q \sigma_p & \sigma_q^2 \\
\end{bmatrix}
\]

Analogously, the measurement vector used by BC-DSSE can be defined as \( z_{\psi \phi} = (V i^r i^p)^T \) and the corresponding covariance matrix is obtained by means of the Jacobian \( \frac{\partial f}{\partial z_{\psi \phi}} \) of the transformation function vector \( g \) with respect to the independent measurements \( z_{\psi \phi} \). With the simplifications introduced in Section III-A the transformation is linear and its Jacobian becomes:

\[
\frac{\partial g}{\partial z_{\psi \phi}^T} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & -1 \\
\end{bmatrix}
\]

and thus (22) follows thoroughly.

**TABLE IV**

<table>
<thead>
<tr>
<th>Correlation level</th>
<th>Improvement on RMSE [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch 25</td>
<td>1.7</td>
</tr>
<tr>
<td>Branch 39</td>
<td>5.7</td>
</tr>
<tr>
<td>Branch 75</td>
<td>5.1</td>
</tr>
</tbody>
</table>

where correlation factors are obtained from historical data or run-time estimations. To highlight the behaviour of the estimation results in presence of such a mismatch, correlation factors used as reference case have been decreased, so that both underestimation and overestimation are possible at high degree. Thus, a correlation factor of 0.6 has been assumed for the intra-node and the inter-node correlation of the loads, while a correlation factor equal to 0.675 has been used for the inter-node correlation of the generators. Starting from this reference case, tests considering all the correlations have been performed, with the correlation factors adopted for the DSSE that range from 1/4 to 5/4 of the reference values in order to simulate the errors in covariance estimation.

Results clearly show that considering correlations into the estimator leads to significant benefits also when correlation factors are not perfectly matched. Fig. 10 reports, as an example, the results for the current estimation in branch 45.
ACKNOWLEDGEMENTS

This work was supported by Regione Autonoma della Sardegna, L.R. 7/2007: “Promozione della ricerca scientifica e dell’innoazione tecnologica in Sardegna, annualità 2012”, CRP-60511.

REFERENCES


Carlo Muscas (M’98) received the M.S. degree in Electrical Engineering in 1994 at the University of Cagliari, Italy, where he was Assistant Professor from 1996 to 2001 an where since 2001 he has been Associate Professor of Electrical and Electronic Measurement. His research activity mainly concerns the study of power quality phenomena, the measurement of synchrophasors and the implementation of distributed measurement systems for modern electric grid. He is author or co-author of more than one hundred scientific papers and is currently Associate Editor for the IEEE Transactions on Instrumentation and Measurement.

Marco Pau (S’14) received the M.S. degree in Electric Engineering from the University of Cagliari, Italy in 2011. He is a Ph.D. Student at the Department of Electrical and Electronic Engineering, University of Cagliari. His research activities focus on measurement systems applied to electrical distribution grids and designed for distribution system state estimation.

Paolo Attilio Pegoraro (S’03-M’06) received the M.S. degree in telecommunications engineering (cum laude) and the Ph.D. degree in electronic and telecommunications engineering from the University of Padova in 2001 and 2005, respectively. Currently, he holds a post-doctoral position with the University of Cagliari where his research activities focus on development of new measurement techniques for modern power networks.

Sara Sulis (S’04, M’06) is assistant professor of Electrical and Electronic Measurements at the University of Cagliari. In December 2013 Dr. Sulis obtained the habilitation to the position of associate professor. Her research activity concerns distributed measurement systems designed to study reliable methodologies to perform both state estimation and harmonic sources estimation of distribution and transmission networks.