



# Exploiting autoencoders for three-phase state estimation in unbalanced distributions grids

P.N. Pereira Barbeiro<sup>a,b</sup>, H. Teixeira<sup>a,b</sup>, J. Krstulovic<sup>a,b</sup>, J. Pereira<sup>a,b</sup>, F.J. Soares<sup>a,b,\*</sup>

<sup>a</sup> INESC TEC—INESC Technology and Science (formerly INESC Porto), Porto, Portugal

<sup>b</sup> Faculty of Engineering, University of Porto, Porto, Portugal

## ARTICLE INFO

### Article history:

Received 24 August 2014

Received in revised form

18 December 2014

Accepted 9 February 2015

Available online 28 February 2015

### Keywords:

Autoencoders

Distribution grids

Neural networks

Three-phase state estimation

Unbalanced loads

Smart metering

## ABSTRACT

The three-phase state estimation algorithms developed for distribution systems (DS) are based on traditional approaches, requiring components modeling and the complete knowledge of grid parameters. These algorithms are capable of dealing with the particular characteristics of DS but cannot be used in cases where grid topology and parameters are unknown, which is the most common situation in existing low voltage grids.

This paper presents a novel three-phase state estimator for DS that enables the explicit estimation of voltage magnitudes and phase angles in all phases, neutral, and ground wires even when grid topology and parameters are unknown. The proposed approach is based on the use of auto-associative neural networks, the autoencoders (AE), which only require an historical database and few quasi-real-time measurements to perform an effective state estimation.

Two test cases were used to evaluate the algorithm's performance: a low and a medium voltage grid. Results show that the algorithm provides accurate results even without information about grid topology and parameters. Several tests were performed to evaluate the best AE configuration. It was found that training an AE for each network feeder leads generally to better results than having a single AE for the entire system. The same happened when different AE were trained for each network phase in comparison with a single AE for the three phases.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

The operation and management of distribution systems (DS) is becoming more complex due to the multitude of assets that are starting to be deployed in these networks. New storage devices, flexible loads, microgeneration units, distributed generation and electric vehicles can be very useful to improve DS efficiency, but only if new operational methods and tools are used by distribution system operators (DSO) to efficiently manage and control all these distributed resources. A tool of unquestionable value for this purpose is a state estimation (SE) algorithm suited for DS. Such tool will aid the DSO to monitor and operate the DS in quasi-real-time, similarly to what already happens in transmission networks.

The majority of the SE techniques existing nowadays are based on conventional methods, being the most common the weighted

least squares (WLS) [1]. These methods were initially designed for transmission networks and their success relies on the complete knowledge of grid technical parameters and topology and a big amount of quasi-real-time measurements available (high redundancy).

Although conventional methods are very accurate for transmission networks, their application to DS is not straightforward since these networks usually have multi-phase lines (sometimes with an asymmetrical cable infrastructure), loads and microgeneration/dispersed generation unevenly distributed among phases and a reduced number of quasi-real-time measurements available. An even more important issue is the lack of information about network topology and parameters, particularly at the low voltage (LV) level.

Some changes to traditional SE methods have been proposed in the literature, which make them capable of coping with some of the DS characteristics. When the number of quasi-real-time measurements is not enough to guarantee observability of a WLS state estimator, pseudo-measurements can be generated using historical data or load curve assessment, as described in [2–6]. The authors of [7,8] developed WLS based methods to perform single-phase

\* Corresponding author at: INESC TEC (formerly INESC Porto), Campus da FEUP, Rua Dr. Roberto Frias, 378, 4200-465 Porto, Portugal. Tel.: +351 22 209 4212; fax: +351 22 209 4050.

E-mail address: [fsoares@inescporto.pt](mailto:fsoares@inescporto.pt) (F.J. Soares).

AE	autoencoders
ANN	artificial neural networks
DMS	distribution management systems
DS	distribution systems
DSO	distribution system operator
DTC	distribution transformer controller
EPSO	evolutionary particle swarm optimization
GPRS	general packet radio service
HLRR	hidden layer reduction rate
LV	low voltage
MSE	mean square error
MV	medium voltage
POCS	projection onto convex sets
SE	state estimation
WLS	weighted least squares

SE assuming a smart grid scenario [7,8]. The contribution of phasor measurements units to enhance distribution SE algorithms has been analyzed in detail in [9–11]. In other works, such as [2,12], authors have applied artificial neural networks (ANN) to estimate system state variables. In all these studies, the unbalanced nature of the DS was never taken into account (balanced load conditions were always assumed). This is a key point since only a three-phase analysis allows getting a complete quasi-real-time snapshot of the system, something indispensable for its correct monitoring and control. As an example, in case of emergency conditions (e.g. over or under voltage profiles), three-phase SE may be used to implement precise load control actions to the costumers that are effectively contributing for the problem instead of affecting all the clients in a given area.

This issue is tackled in several publications available in the literature that specifically address the three-phase state estimation topic [13–19]. Two common points can be found in these works: neutral voltage is never explicitly estimated since Kron's reduction is usually applied; and the algorithms are based on complex three-phase mathematical equations that require the total knowledge of grid technical parameters and topology. As in the majority of these works phase-to-neutral voltage is disregarded, the voltage imbalance is often miscalculated. This inaccuracy is particularly important in 4-wire LV grids, where connection of single-phase loads provokes imbalances in the system and the appearance of a return current divided by the neutral and ground circuits. This results in a reduction of phase-to-neutral voltage for single-phase customers. The importance of knowing voltage and neutral currents becomes evident when observing the more common neutral designs in DS, which are normally favorable to the appearance of neutral and ground currents (due to the unbalanced nature of the loads). Knowing neutral voltages has also a great importance when its use is related with power quality, safety or energy losses. Taking into account the size and complexity of some DS, it is reasonable to admit that a three-phase state estimation tool based on the conventional approaches will increase the computational burden in a considerable manner. This would probably make these techniques unfeasible for quasi-real-time applications.

An effective distribution SE algorithm should therefore be capable of dealing with all the particular characteristics of DS, take advantage of the telemetry measurements that may be available (both in quasi-real-time or historical records stored in distribution management systems—DMS) and, at the same time, be fast enough to run in quasi-real-time.

This paper proposes an innovative method for complete three-phase state estimation (voltage magnitudes and phase angles in all phases, neutral and ground wire), which is based on the use of a

particular kind of ANN—the autoencoders (AE). The method takes advantage of the data gathered from the smart metering infrastructure, both in quasi-real-time and historical, and then uses artificial intelligence capabilities to learn the behavior of the grid and accurately estimate the state of the system at any moment. This avoids two vital steps of the traditional state estimation algorithms: (1) modeling the complex three-phase equations, which may lead to heavy iterative algebraic calculations and numerical/convergence problems and (2) characterizing all the grid parameters, which are often unknown in LV grids. Another important feature is that the proposed method is very flexible regarding the type of electrical variables that can be passed to the SE algorithm, meaning that a full exploitation of all the available telemetry information is always performed. In other words, measurements such as active/reactive powers consumed or generated, active/reactive power flows, current magnitudes, voltage magnitudes/phase angles, power factors, and even energy measurements can be used whenever they are available in quasi-real-time. For all these reasons, the proposed method can be used for any grid, independently of the neutral and earthing system adopted type of loads/generators present in the system and quasi-real-time measurements available.

The algorithm performance was tested in two typical Portuguese distribution grids: a 4-wire LV grid with several single and three-phase loads, microgeneration units and smart meters and a 3-wire medium voltage (MV) grid with distributed generation, residential and industrial clients and different telemetry technologies. Accuracy and running times were analyzed under different conditions. The tests performed included: varying the number of neurons in the AE hidden layer, different AE types (global vs. local) and different scenarios regarding number and type of technologies capable of transmitting measurements in quasi-real-time. For the purpose of this work, the term “quasi-real-time” is used in the sense of measuring the variables in a short period of time, around 15 min (or even less, depending on the communication infrastructure).

## 2. Autoencoders applied to the three-phase state estimation problem

AE are frequently applied in areas related with pattern recognition and reconstruction of missing sensor signals [20,21]. Their application in the power systems area is confined to a few works. In two interesting papers, [22] and [23], the authors used AE to reconstruct missing measurements in the SCADA of the DMS, to identify errors in breaker status and to find the power system topology. Nevertheless, the use of AE as “the core” of a three-phase state estimation algorithm was never implemented nor tested. AE, or auto-associative neural networks, are feedforward neural networks where the size of the output layer (number of neurons) is always equal to the input layer. The typical architecture of an AE is a neural network with only one middle layer (Fig. 1). This simple architecture is frequently adopted since in the generality of the applications, networks with more hidden layers does not bring any benefit and have proved to be more difficult to train [24].

With adequate training, an AE learns the data set pattern and stores in its weights manifold information about the training data. The training process of an AE is conducted to display the output equal to its input. The first half of the AE approximates the function  $f$  that encode the input space to the space compressed  $S'$ . For an input space composed by an  $n$ -dimensional input vector  $x_k$  ( $k = 1, 2, \dots, n$ ), the output values of the hidden neurons form a  $m$ -vector given by:

$$h_p = f_a(W_{a(m,n)}x_k + B_{am}), \quad k = 1, 2, \dots, n \text{ and } p = 1, \dots, m, \quad (1)$$

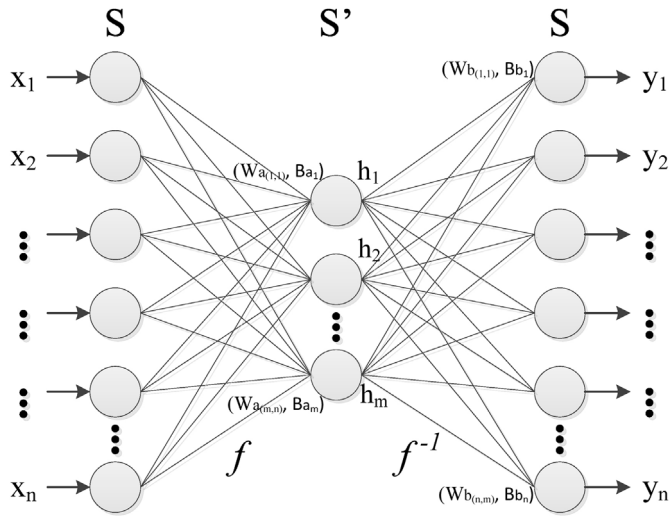


Fig. 1. Autoencoder architecture.

where  $W_a$  is the input-to-hidden  $m \times n$  weight matrix and  $B_a$  is the hidden neurons  $m$ -vector of biases.  $f_a$  represents the hidden neurons activation function.

The second half approximates the inverse function  $f^{-1}$  that projects back the set of values in space  $S'$  to the original space  $S$ . This mapping happens through a similar transformation and the output values form again an  $n$ -vector given by;

$$y_k = f_b (W_{b(n,m)} h_p + B_{b_n}), \quad k = 1, 2, \dots, n \text{ and } p = 1, \dots, m, \quad (2)$$

where  $W_b$  is the hidden-to-output  $n \times m$  weight matrix and  $B_b$  is the output neurons  $n$ -vector of biases.  $f_b$  represents the output neurons activation function. For a historical database with  $s$  samples, the optimal weight matrices  $W_a$ ,  $W_b$  and bias vectors  $B_a$ ,  $B_b$  can be found minimizing the mean squared error during the training procedure, according to:

$$J = \frac{1}{s} \sum_{s=1}^s \sum_{k=1}^n (x_k - y_k)^2 \quad (3)$$

Once the AE is trained, if an incomplete pattern is presented, the missing components may be replaced by random values producing a significant mismatch between input and output. Three different approaches can be followed in order to reach convergence and find the missing signals. The approach called projection onto convex sets (POCS) consists basically in iteratively reintroducing the output value in the input, [25]. This convergence method will lead to a value that minimizes the input–output error. It uses alternating linear projections in the input/output space  $S$  to converge to the missing signals. The two other approaches are based on unconstrained and constrained search techniques, where an optimization algorithm is used to discover the missing signals. The search is conducted in a way that minimizes the input–output error. Any of these optimization procedures may be used, but according to some published works [22,23], constrained search appears to yield better results.

The aforementioned concepts were adapted in this paper and applied to the three-phase state-estimation problem. The resulting AE is illustrated in Fig. 2. A constrained search approach was used to estimate the system state variables: voltage magnitude and phase angle in all the buses/phases that are not being monitored in quasi-real-time.

### 3. Developing the distribution state estimator

The methodology followed to develop the distribution SE comprises three main steps:

1. Building the historical data set;
2. Training the AE (offline procedure);
3. Building the distribution SE algorithm.

These steps are described in detail in the following sections.

#### 3.1. Building the historical data set

In future MV grids, synchronized measurements gathered by telemetry equipment, such as distribution transformer controllers (DTC), will be stored in a database of the DMS. The stored measurements can then be used as the historical data set of the SE. A similar situation is foreseen for LV grids, but in this case data will be gathered by smart meters.

There is no a priori indication of the quantity of data that the historical data set should contain. The AE learning process improves with the amount of data available, but only until a certain point.

As load demand usually obeys to well-defined seasonal patterns, the historical data set can include several of these patterns. In these situations, an AE per pattern should be used in order to achieve better results. Alternatively, a single AE can be trained for the entire historical data set, but results will not be as accurate. Either way, the AE should be trained with an enough number of samples in order to yield good results. The same stands for different grid topologies. If the historical data set includes data of different grid topologies, several AE should be trained (one per grid topology). Again, it is possible to train a single AE for more than one topology, yet results will be less accurate.

#### 3.2. Training the AE

The flowchart of the algorithm used to train the AE is presented in Fig. 3.

After defining the historical data set, a standardization procedure is run with the goal of pre-treating the input and output training data set. This is a simple scale adjustment process, where the range of the input and output values are adjusted to be contained in the interval  $[-1, 1]$ . This procedure fits the input variables to the range of the activation function. The method selected to perform the standardization was the “Min–Max”, expressed in Eq. (4). This is the best standardization procedure for cases where the minimum and maximum values of the data set are known.

$$y' = \frac{y - \min_a}{\max_a - \min_a} \times (\max_A - \min_A) + \min_A \quad (4)$$

where  $y'$  is the standardized value for the considered variable;  $y$  is the variable value in the original representation interval;  $\min_a$  is the minimum value of the “original” range of values;  $\max_a$  is the maximum value of the “original” range of values;  $\min_A$  is the minimum value of the standardized range of values ( $-1$ );  $\max_A$  is the maximum value of the standardized range of values ( $1$ ).

The next is selecting a type of AE. The “global AE” refers to the utilization of a single AE for the entire network. “Local AE” is referred to the use of an AE per network feeder. For both cases, the algorithm tests two more possibilities: a single AE for the three phases and an AE per phase.

Afterwards, the parameters to be used during the training stage are defined. Some of them are typical values, while others are case dependent (influenced by the characteristics of the problem, type of network, etc.) and require a trial and error approach to be properly tuned. As an example, there is no a priori indication of

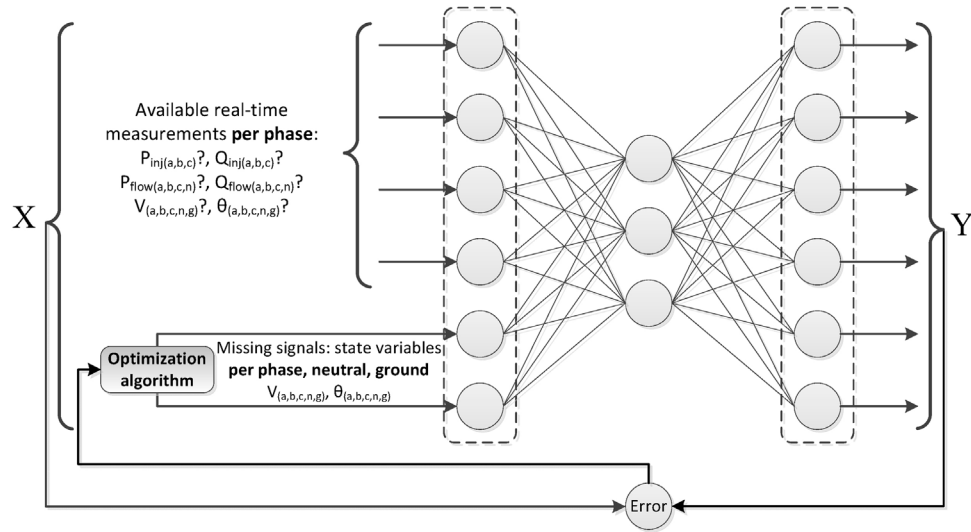


Fig. 2. Autoencoder applied to the three-phase state estimation problem with the constrained search approach.

an adequate hidden layer reduction rate (measured as the ratio between the number of neurons in the middle layer and in the input layer of the AE). For this reason, tests over several parameters of the AE were performed beforehand to define the best architecture

for the three-phase state estimation problem. For the hidden layer, for example, different numbers of neurons were tested. A relatively high initial value for the hidden layer reduction rate was assumed at first, being gradually reduced in a stepwise manner while the AE results were improving. The optimal hidden layer reduction rate was found when the AE results began to worsen.

A resilient back-propagation algorithm was used to train the AE. It belongs to the most widely used class of algorithms for supervised learning of neural networks. It works as the name suggests: after propagating an input through the AE, the error is calculated and then it is propagated back through the network while weights are adjusted in order to make the error smaller. An interesting particularity of this algorithm is that instead of training combined data, the training data set is executed sequentially, one input at a time, minimizing the MSE for the entire training data set. This also avoids getting stuck in a local minimum. Experimental training tests were carried out to select the most appropriate activation function for the hidden and output layers. These activation functions can be modeled by different mathematical functions, being the most common the threshold, the sigmoid and step wise function. For the specific problem under analysis, results have shown that the AE performs better when non-linear activation functions are used in both layers. Thus, a symmetric sigmoid was adopted as the activation function for both the hidden and the output layers. This function can be represented by:

$$\begin{cases} f(v) = \frac{1}{1 + e^{-cv}} \\ -1 \leq f(v) \leq 1 \end{cases} \quad (5)$$

where  $f(v)$  is the output of the respective neuron,  $v$  is the sum of all inputs (affected by the respective weights) plus the corresponding bias of that neuron and the constant  $c$  is the slope parameter of the sigmoid function.

The number of training epochs is another important parameter to fine-tune the weights of the AE. An overstated number can lead to overfitting, while the opposite is very likely to lead to underfitting. This effect can be overcome by analyzing the evolution of the Mean Square Error (MSE) of the test set. For the purpose of this work, the training algorithm was set to perform 100 epochs at a time and, in the end, analyze the MSE of the test set. If it decreases, the algorithm runs 100 epochs more, otherwise it stops.

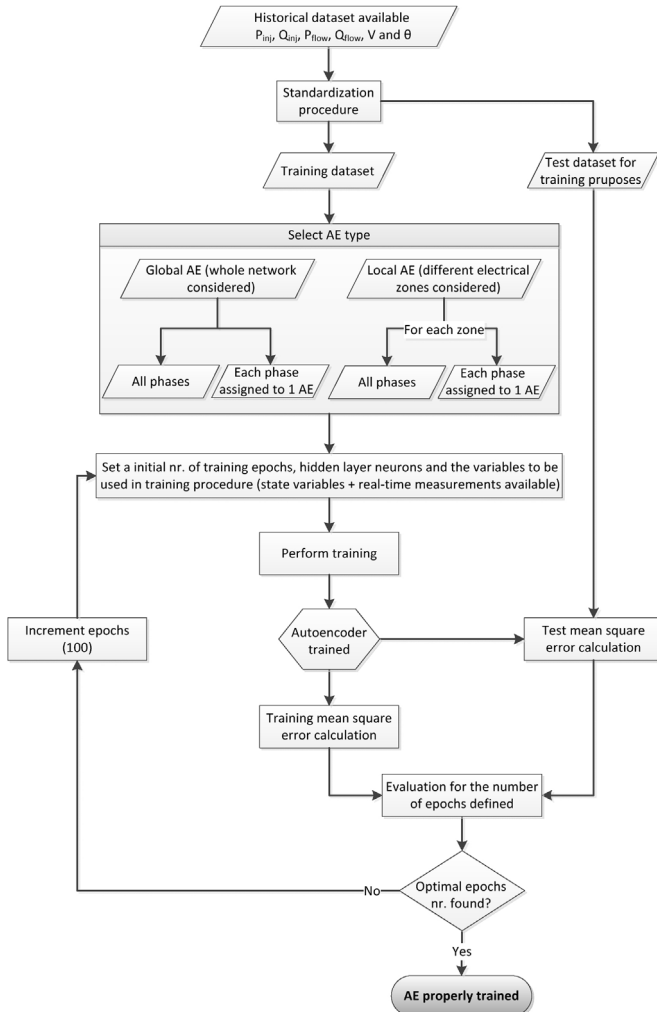


Fig. 3. Flowchart of the algorithm used to train the AE.

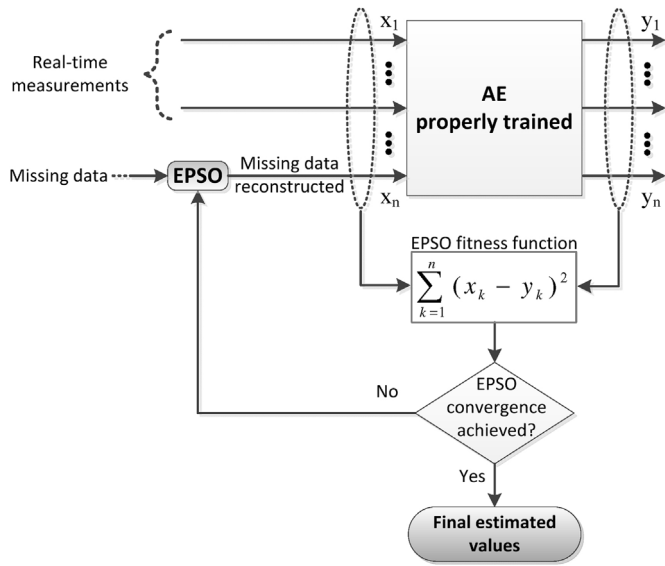


Fig. 4. Scheme of the distribution state estimation algorithm.

### 3.3. Building the distribution SE algorithm

The distribution SE algorithm consists basically in adding an optimization procedure for reconstructing the missing variables to the AE already trained (see flowchart of Fig. 4). The meta-heuristic selected to reconstruct the missing signals was an Evolutionary Particle Swarm Optimization (EPSo), which has been successfully applied in several problems in the power systems area [26–28]. The fitness function of the EPSo was defined to minimize the square error between the AE input and output, given by:

$$\varphi(\varepsilon) = \sum_{k=1}^n (x_k - y_k)^2 \quad (6)$$

After having the AE properly trained and the EPSo incorporated to reconstruct the missing variables, its performance is evaluated by estimating the state of the networks presented in Section 4.

Two parameters were used to access the AE performance: running time and accuracy. Results obtained for both parameters are presented in Section 5.

## 4. Case studies and AE performance evaluation

### 4.1. Low voltage network

A typical Portuguese LV network was used as the first test case (Fig. 5). The secondary substation has a 15 kV/400 V/230 V transformer with a rated power of 100 kVA. The transformer has a delta connected winding on its primary and a wye/star connected winding on its secondary with accessible neutral (grounded through a reactance at the substation level). This is a popular configuration in Portugal and widely used in several other European countries. This configuration leads to the existence of returning currents in the neutral wire whenever load imbalances exist in the network.

The grid contains 57 consumers with contracted powers that vary between 3.45 and 6.9 kVA for single-phase and between 6.9 and 13.8 kVA for three-phase clients. Since a significant amount of single-phase loads is present, load is not distributed equally among phases, generating imbalances between them. The network has a total of 33 nodes and a peak load of 92.3 kW. Fig. 5 shows the network schematics. Due to space restrictions, only the three main feeders were represented. Details about the grid connection of a single-phase and three-phase client are also represented.

Several microgeneration units (photovoltaic panels) were added and randomly distributed through the network clients, totalizing ~25% of the secondary substation transformer capacity (ca. 25 kVA). The microgeneration units were assumed to be single-phase. Whenever a consumer is a three-phase type and has a photovoltaic panel, this unit is randomly connected to one of the phases. Table 1 shows the existing consumers and the microgeneration units connected per phase.

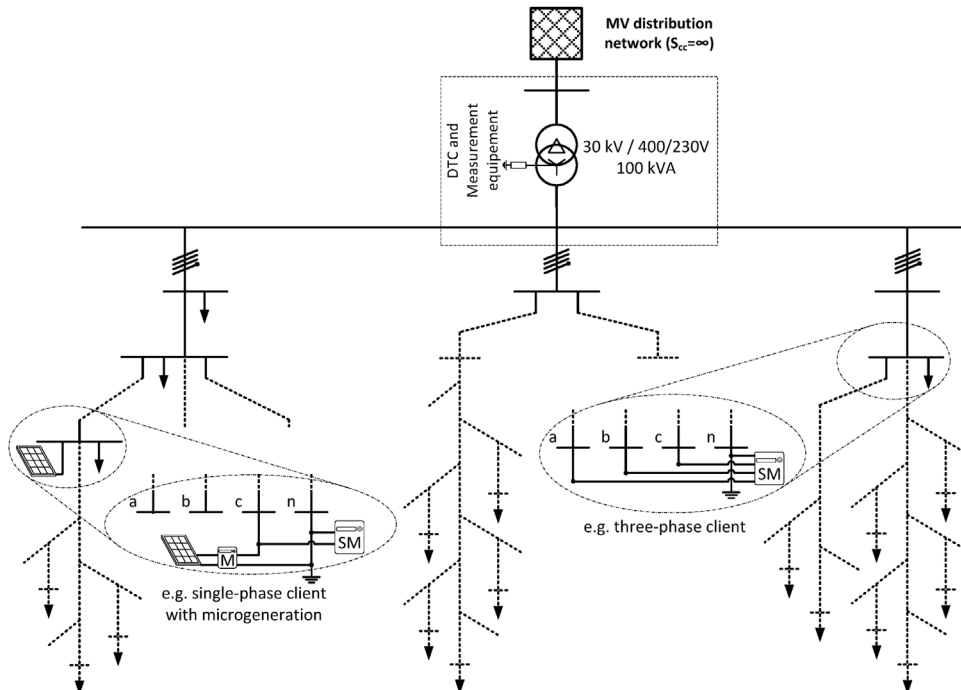


Fig. 5. LV network schematics.



**Table 1**  
Consumers and microgeneration distribution.

Client type (connection phase)	Contracted power	Microgeneration installed capacity
<b>Single-phase (a)</b>	1.15 kVA (bus 6) 3.45 kVA (buses 2, 8, 10, 11, 12, 19, 21, 22, 24, 29) 3 clients with 3.45 kVA (bus 7) 6.9 kVA (buses 9, 13, 22, 27) 13.8 kVA (bus 17)	3.45 kVA (buses 9, 13, 27) 5.75 kVA (bus 17)
<b>Single-phase (b)</b>	3.45 kVA (buses 2, 8, 9, 12, 13, 18, 19, 20, 21, 22, 25, 27, 29, 32) 6.9 kVA (buses 16, 31) 13.8 kVA (buses 10, 26)	3.45 kVA (buses 16, 31) 5.75 kVA (buses 10, 26)
<b>Single-phase (c)</b>	3.45 kVA (buses 5, 8, 9, 13, 18, 20, 22, 23, 24, 27, 29, 32, 33) 6.9 kVA (buses 11, 29) 13.8 kVA (buses 25, 30)	3.45 kVA (buses 11, 29) 5.75 kVA (buses 25, 30)
<b>Three-phase</b>	3 clients with 17.25 kVA (bus 24)	3 clients with 5.75 kVA (bus 24)

#### 4.2. Medium voltage network

The second test case is a MV network that has a transformer of 10 MVA in the primary substation, with a wye/star connected winding on its primary side (high voltage level) and delta connected winding on its secondary (MV level). A total of 115 secondary substations exist in the grid, being 95 from residential low voltage grids and 20 from industrial clients. All MV/LV transformers are three-phase. The network has a total of 303 nodes and a peak load of 5.8 MW. Fig. 6 shows the network schematics.

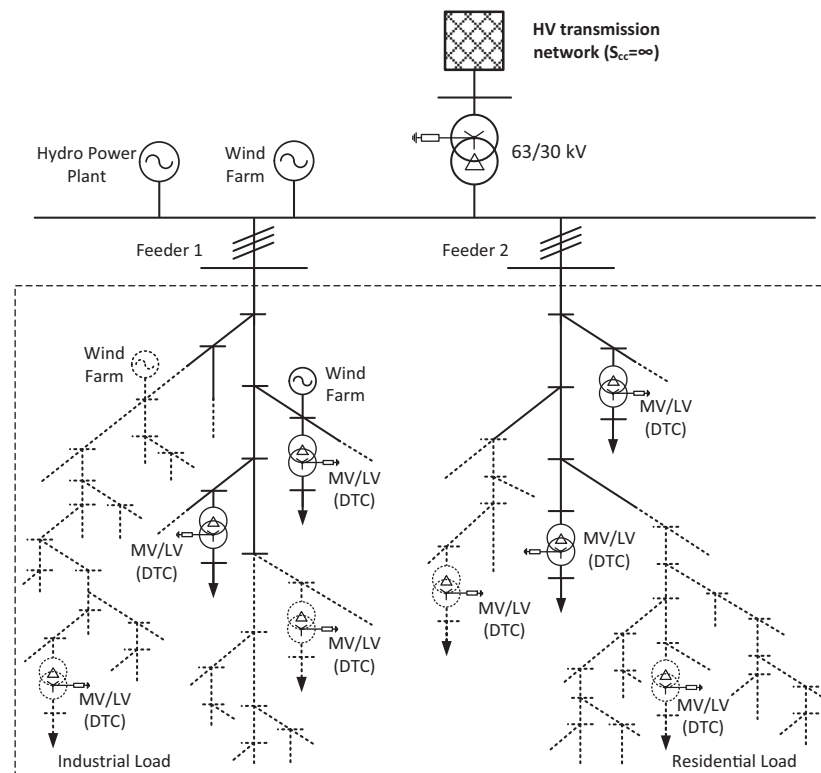
#### 4.3. Modelling load and microgeneration variability

Regarding load, for both LV and MV networks, the only data available was the monthly average of the aggregated load at the secondary and primary substation levels, respectively. In order to

represent consumers with different behaviors, distinct load diagrams were generated scaling down the average diagram for each client and by adding Gaussian noise (average of 0% and standard deviation of 10%) to it.

For the photovoltaic panels in the LV network, 5 real generation diagrams were used. These 5 profiles represent different days (e.g. sunny, cloudy, rainy, etc.) and were obtained from a real meteorological station [29]. They were randomly distributed by the existing units according to their probability of occurrence in a typical Portuguese summer (for the LV grid) and winter (for the MV grid). A similar method was used for the distributed generation units (wind farms and hydro plants) of the MV network.

The high variability introduced by the different load/generation diagrams can be regarded as a worst case scenario to test the AE performance. It is very likely that if a real historical database was used, the variability would not be as high and the AE would perform significantly better.



**Fig. 6.** MV network schematics.

**Table 2**  
Quasi-real-time measurements scenarios.

Network	Scenario	Nr. of real time measurements (m)	State variables to be estimated (n)	m/n
<b>MV</b>	0	18	1464	0.012
	1	1368	792	1.727
<b>LV</b>	0	30	216	0.139
	1	110	184	0.598

#### 4.4. Scenarios of existing telemetry equipment

In order to emulate a smart grid environment, some additional features and equipment were assumed to exist in these networks. In the primary substation of the MV network telemetry equipment (such as phasor measurement units or remote terminal units) was assumed to exist with the capability of monitoring in quasi-real-time the following variables: active and reactive power flows in the transformer and in the MV feeders and voltage (magnitude and phase) at the high and medium voltage side of the transformer. The secondary substations were assumed to house a DTC, as well as the associated measurement equipment, with the capability of monitoring the same variables as in the primary substation.

It was also considered that each LV customer has a smart meter to monitor his consumption and communicate it to the DTC for billing purposes. The customers that own a microgeneration unit have an additional smart meter for measuring its power production. As it happens in some real smart grid test sites, not all smart meters are capable of transmitting data in quasi-real time due to communication infrastructure restrictions. Only some, which use, for instance, *general packet radio service* (GPRS) technology, have that capability. It was assumed that their active power ( $P$ ), reactive power ( $Q$ ) and voltage magnitude ( $V$ ) measurements are synchronized. Phase angles were assumed not to be measured as the majority of the smart meters foreseen to be deployed in LV grids do not have this capability. In terms of data accuracy, smart meters are usually categorized in classes, according to the confidence level specified by the manufacturers. After checking the technical specification of a large set of smart meters available in the market, some typical values were assumed. Voltage measurements were considered with  $\pm 1\%$  accuracy and  $P$  and  $Q$  measurements with  $\pm 2\%$ , all with a confidence level of 95%.

Taking into account the referred assumptions, two scenarios of equipment capable of transmitting data in quasi-real-time were considered. They are presented in Table 2.

In scenario 0, for the MV grid, only quasi-real-time measurements from equipment existing in the primary substation were assumed to exist. Similarly, only measurements from the secondary substation were assumed to exist in the LV grid. Scenario 1 is more optimistic. For the MV grid, it was assumed that all the secondary substations have a DTC with the capabilities referred before. For the LV grid, smart meters with the capability of transmitting data in quasi-real-time were assumed to exist in the load buses farthest from the secondary substation.

#### 4.5. Historical data generation

As there is no historical data for both networks, a three-phase power flow algorithm based on backward/forward sweep method [30] was used to generate it in time steps of 15 minutes (considering a single grid topology). Twelve weeks of fictitious data were generated according to the information provided in the two previous subsections. The historical database was divided in three sets: 8 weeks were used for the training set, 1 for the test set (used during the AE training) and the remaining to the evaluation set. Results regarding the AE performance (presented in the following section) are referred to the evaluation set.

### 5. Results

#### 5.1. Low voltage network

The accuracy of the AE performance was evaluated using the MAE, as shown in Eq. (1), where  $y_i$  represents the real values (generated with a load flow),  $f_i$  the AE estimates and  $n$  the number of samples.

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| \quad (7)$$

Table 3 summarizes the MAE and the maximum absolute error, in p.u., obtained during the tests performed to define the optimal number of neurons in the hidden layer. These tests were only conducted for the global AE for the three phases. In the first test, a

**Table 3**  
Results of the HLRR tests.

	HLRR 0.6–148 neurons on hidden layer				HLRR 0.1–25 neurons on hidden layer			
	Scenario 0		Scenario 1		Scenario 0		Scenario 1	
	MAE	Max.	MAE	Max.	MAE	Max.	MAE	Max.
<b>Phase a</b>	0.0072	0.0682	0.0040	0.0282	0.0041	0.0301	0.0015	0.0133
<b>Phase b</b>	0.0082	0.0377	0.0056	0.0364	0.0048	0.0250	0.0023	0.0249
<b>Phase c</b>	0.0095	0.0606	0.0059	0.0480	0.0051	0.0358	0.0025	0.0261
<b>Neutral</b>	0.0052	0.0501	0.0029	0.0241	0.0037	0.0339	0.0014	0.0129
<b>Avg.</b>	0.0075	–	0.0046	–	0.0044	–	0.0019	–
<b>Max.</b>	–	0.0682	–	0.0480	–	0.0358	–	0.0261

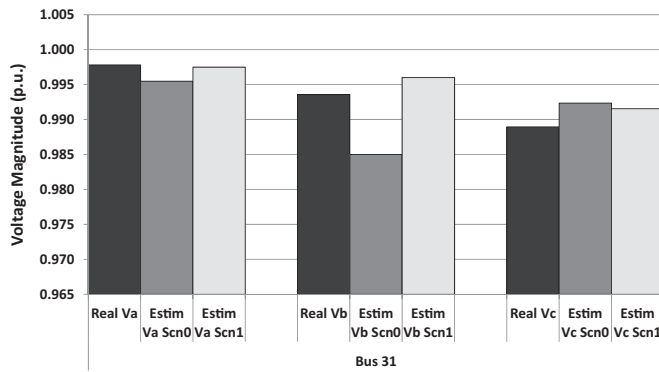
**Table 4**  
Results for the types of AE tested—LV grid, scenario 0.

	Global AE 3 phase		Global AE per phase		Local AE 3 phase		Local AE per phase	
	MAE	Max.	MAE	Max.	MAE	Max.	MAE	Max.
<b>Phase a</b>	0.0041	0.0301	0.0028	0.0323	0.0033	0.0245	0.0016	0.0224
<b>Phase b</b>	0.0048	0.0250	0.0034	0.0249	0.0034	0.0230	0.0018	0.0179
<b>Phase c</b>	0.0051	0.0358	0.0035	0.0272	0.0038	0.0352	0.0019	0.0183
<b>Neutral</b>	0.0037	0.0339	0.0035	0.0369	0.0025	0.0184	0.0032	0.0383
<b>Avg.</b>	0.0044	–	0.0033	–	0.0032	–	0.0021	–
<b>Max.</b>	–	0.0358	–	0.0369	–	0.0352	–	0.0383

**Table 5**

Results for the types of AE tested—LV grid, scenario 1.

	Global AE 3 phase		Global AE per phase		Local AE 3 phase		Local AE per phase	
	MAE	Max.	MAE	Max.	MAE	Max.	MAE	Max.
Phase a	0.0015	0.0133	0.0009	0.0057	0.0015	0.0085	0.0004	0.0036
Phase b	0.0023	0.0249	0.0014	0.0138	0.0017	0.0120	0.0006	0.0069
Phase c	0.0025	0.0261	0.0010	0.0106	0.0016	0.0126	0.0004	0.0053
Neutral	0.0014	0.0129	0.0009	0.0075	0.0013	0.0099	0.0007	0.0075
Avg.	0.0019	–	0.0011	–	0.0015	–	0.0005	–
Max.	–	0.0261	–	0.0138	–	0.0126	–	0.0075

**Fig. 7.** Voltage magnitude for the bus with the highest maximum absolute error—global AE for three phases.

hidden layer reduction rate (HLRR) of 0.6 was considered. Afterwards, this value was gradually reduced in steps of 0.05 while the MAE was reducing. The optimum HLRR was found when the MAE began to increase. The optimal HLRR was found to be 0.1, corresponding to 25 neurons. A HLRR of 0.1, instead of 0.6, produced more accurate results in both scenarios analyzed.

The results obtained for the various types of AE tested are presented in [Tables 4 and 5](#), respectively, for scenario 0 and 1. The 'local AE per phase' generated the lowest average MAE, while the 'local AE for three phases' generated the lowest maximum absolute error. After a careful analysis of the results, it was verified that for all the AE the probability of occurring a maximum absolute error higher

than 0.02 p.u. was always lower than 0.5%. This value reveals the good performance achieved by the AE.

Given the large number simulations performed and results compiled, only some examples were selected to be presented in graphical format. [Fig. 7](#) shows the real and estimated voltage magnitude in the three loaded phases for the bus with the highest maximum absolute error recorded during the simulations (global AE for three phases). This figure evidences the unbalanced nature of the system as it shows the voltage difference amid phases.

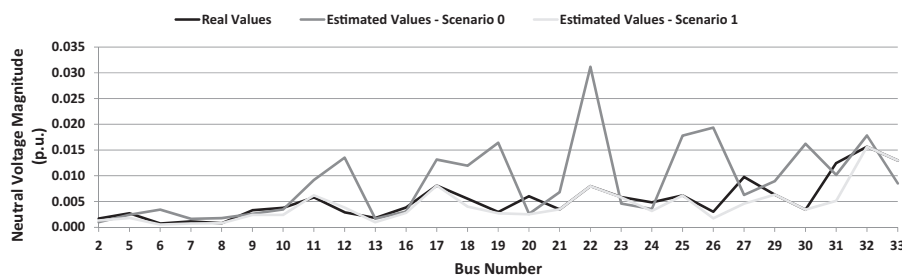
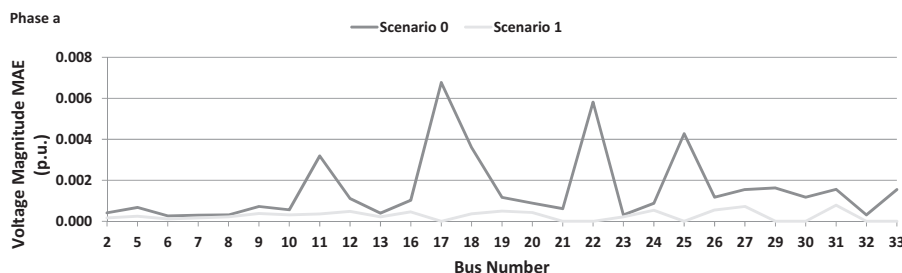
[Fig. 8](#) shows the voltage magnitude in the neutral obtained with the 'global AE for three phases'. It is clearly perceptible that the AE provides better estimates in scenario 1. This is due to the existence of a higher number of smart meters with the capability of transmitting data in quasi-real time.

[Fig. 9](#) depicts the voltage magnitude MAE for phase *a* in all the network buses obtained with the 'local AE per phase'. Again, estimated values are more accurate when more quasi-real-time measurements are considered. For both scenarios, the worst estimated values belong to buses that have both loads and micro-generation connected in the same phase. This is an expected result due to the higher variability of the power injected in these buses.

[Fig. 10](#) shows the performance of the different types of AE used in scenario 1. In general terms, the 'global AE per phase' and the 'local AE per phase' produce more accurate results than the others.

## 5.2. Medium voltage network

The results obtained for the MV follow the same trend of those of the LV network. They were compiled in [Tables 6 and 7](#), respectively,

**Fig. 8.** Voltage magnitude in the neutral—global AE for three phases.**Fig. 9.** Voltage magnitude MAE for phase *a*—local AE per phase.



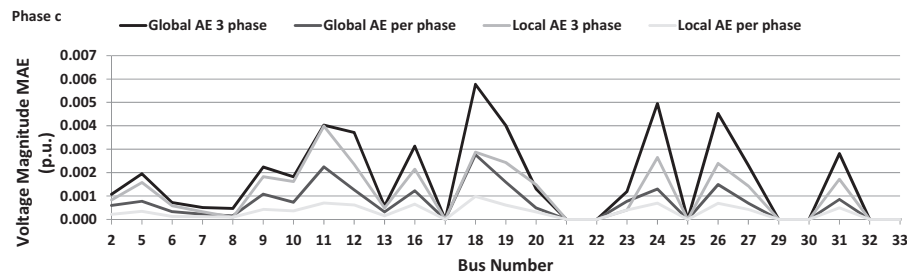


Fig. 10. Performance of the different types of AE used—scenario 1.

Table 6

Results for the types of AE tested—MV grid, scenario 0.

	Global AE 3 phase		Global AE per phase		Local AE 3 phase		Local AE per phase	
	MAE	Max.	MAE	Max.	MAE	Max.	MAE	Max.
Phase a	0.0108	0.0478	0.0104	0.0340	0.0110	0.0460	0.0104	0.0345
Phase b	0.0143	0.0708	0.0143	0.0518	0.0144	0.0588	0.0142	0.0460
Phase c	0.0154	0.0711	0.0155	0.0576	0.0156	0.0641	0.0151	0.0508
Avg.	0.0135	–	0.0134	–	0.0136	–	0.0133	–
Max.	–	0.0711	–	0.0576	–	0.0641	–	0.0508

Table 7

Results for the types of AE tested—MV grid, scenario 1.

	Global AE 3 phase		Global AE per phase		Local AE 3 phase		Local AE per phase	
	MAE	Max.	MAE	Max.	MAE	Max.	MAE	Max.
Phase a	0.0041	0.0296	0.0031	0.0115	0.0040	0.0175	0.0013	0.0099
Phase b	0.0055	0.0399	0.0035	0.0178	0.0051	0.0253	0.0017	0.0089
Phase c	0.0059	0.0430	0.0039	0.0160	0.0055	0.0255	0.0024	0.0095
Avg.	0.0052	–	0.0035	–	0.0048	–	0.0018	–
Max.	–	0.0430	–	0.0178	–	0.0255	–	0.0099

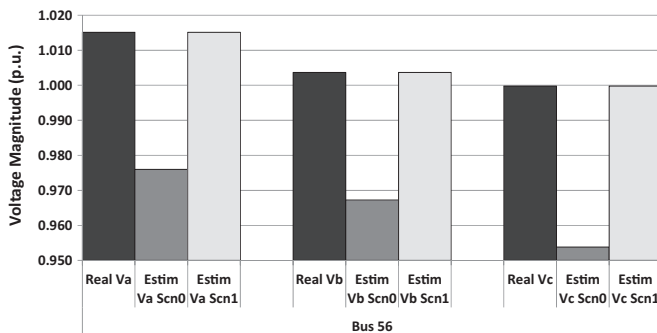


Fig. 11. Voltage magnitude for the bus with the highest maximum absolute error—global AE for three phases.

for scenario 0 and 1. As in the previous case, the 'local AE per phase' was the one that generated the lowest average MAE. This AE was also the one that generated the lowest maximum absolute error.

Fig. 11 shows the real and estimated voltage magnitude in the three loaded phases for the bus with the highest maximum absolute error recorded during the simulations (global AE for three phases).

Fig. 12 shows the voltage magnitude MAE for phase *a* in all the network buses obtained with the 'global AE for three phases'. The buses where the MAE is zero are those who have equipment transmitting data in quasi-real-time.

Fig. 13 shows the voltage angle MAE for the same phase as in the previous figure with the 'global AE for three phases'.

Fig. 14 shows the performance of the different types of AE used in scenario 1. Generally, the 'global AE per phase' and the 'local AE per phase' produce more accurate results than the others.

### 5.3. Running times

The running times of the tested AE are presented in Tables 8 and 9, respectively, for the LV and MV networks. Local AE have a significant advantage in comparison with global ones since parallel processing can be used to greatly reduce running times.

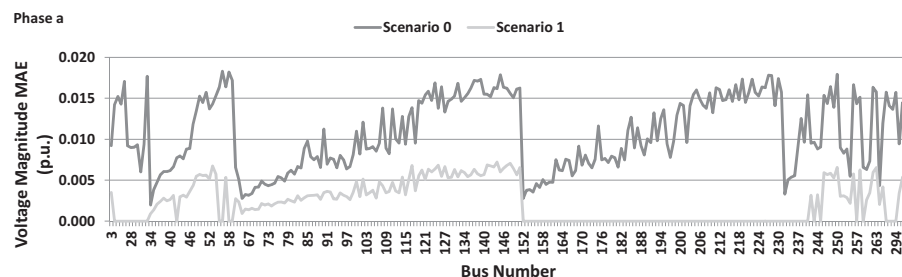


Fig. 12. Voltage magnitude MAE for phase *a*—global AE for three phases.

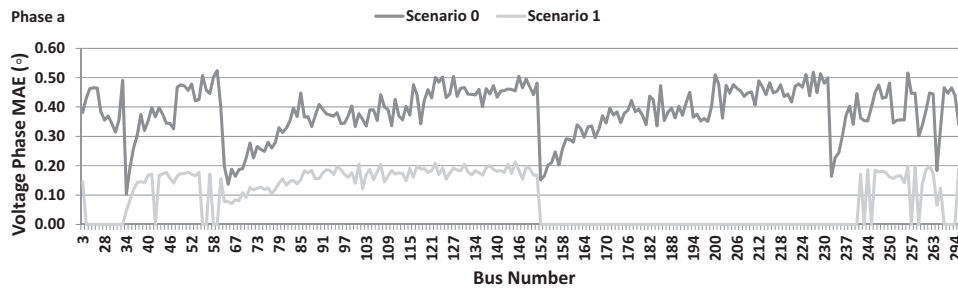


Fig. 13. Voltage angle MAE for phase a—global AE for three phases.

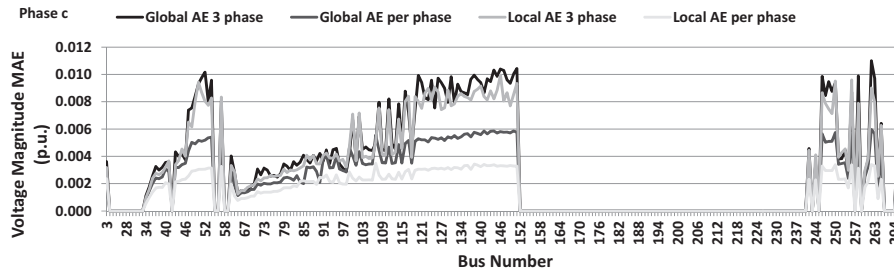


Fig. 14. Performance of the different types of AE used—scenario 1.

Table 8

Running times (in s) for the types of AE tested—LV grid.

Scenario	Global AE 3 phase		Global AE per phase	Local AE 3 phase	Local AE per phase
	HLRR = 0.6	HLRR = 0.1			
0	8.262	1.680	0.250	0.478	0.088
1	10.743	2.120	0.257	0.561	0.145

Table 9

Running times (in s) for the types of AE tested—MV grid.

Scenario	Global AE 3 phase	Global AE per phase	Local AE 3 phase	Local AE per phase
0	10.560	1.860	3.703	0.903
1	31.980	2.620	5.838	1.169

The running times presented were obtained for AE coded in C/C++ programming language and run in a computer with an Intel Core i7-2600 CPU at 3.40 GHz and 8.00 GB of RAM memory. It should be noted that running times are highly influenced by several parameters, such as the convergence criterion of the EPSO or the number of neurons in the AE hidden layer. In fact, a clear trade-off exists between these parameters and the AE accuracy. Thus, it is possible to reduce running times to the detriment of results accuracy (or vice versa). The running times presented in this section should not be seen as absolute values since the algorithms can be coded in more efficient programming languages, leading to lower computation times.

## 6. Conclusions

This paper presented a novel three-phase SE method based on the use of AE. The results obtained show that the proposed approach may be a valid alternative to classic methods if an adequate historical database and some quasi-real-time measurements are available. In fact, when there is no information about the grid parameters and topology, the use of a distribution SE as the one proposed could be the only alternative available to DSO to have a quasi-real-time snapshot of the system.

In order to achieve accurate results, the AE parameters should be carefully tuned during the training process and a robust historical

data set with information about consumption and generation must be gathered.

Several tests were performed to evaluate the best type of AE to use in three-phase SE problems. It was found that an AE trained for each network feeder yielded generally better results than a single AE for the entire system. The same happened when different AE were trained for each network phase in comparison with a single AE for the three phases. In fact, these results are not surprising since it is known that AE performance improves as the number of input variables is lower and the relation between them stronger. However, depending on the characteristics of the system under analysis (i.e. quantity and location of telemetry equipment), the utilization of a different type of AE could be more appropriate.

Results obtained for the running times show that the algorithm is suitable for online applications. It was also found that having “smaller” AE (i.e. an AE per feeder and per phase) instead of a global one leads to lower running times. If parallel processing is used, as it was in this work, this difference might be very significant.

## Acknowledgements

This work was made in the framework of the BEST CASE project (“NORTE-07-0124-FEDER-000056”) financed by the North Portugal Regional Operational Programme (ON.2-O Novo Norte), under the National Strategic Reference Framework (NSRF), through the

European Regional Development Fund (ERDF), and by national funds, through the Foundation for Science and Technology (FCT). It was also co-financed by the COMPETE Programme and the FCT within projects “SMAGIS-PTDC/SEN-ENR/113094/2009” and “DYMONDS-CMU-PT/SIA/0043/2009”.

## References

- [1] A. Abur, A.G. Expósito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, New York, NY, 2004.
- [2] A. Bernieri, G. Betta, C. Liguori, A. Losi, Neural networks and pseudo-measurements for real-time monitoring of distribution systems, *IEEE Trans. Instrum. Meas.* 45 (1996) 645–650.
- [3] E. Manitsas, R. Singh, B. Pal, G. Strbac, Modelling of pseudo-measurements for distribution system state estimation, in: IET-CIRED Seminar 2008: SmartGrids for Distribution, 2008, pp. 1–4.
- [4] R. Singh, B.C. Pal, R.A. Jabr, Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement, *IET Gener. Transm. Distrib.* 4 (2010) 50–59.
- [5] E. Manitsas, R. Singh, B.C. Pal, G. Strbac, Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling, *IEEE Trans. Power Syst.* 27 (2012) 1888–1896.
- [6] V. Miranda, J. Pereira, J.T. Saraiva, Load allocation in DMS with a fuzzy state estimator, *IEEE Trans. Power Syst.* 15 (2000) 529–534.
- [7] G.N. Korres, N.D. Hatziaargyriou, P.J. Katsikas, State estimation in multi-microgrids, *Eur. Trans. Electr. Power* 21 (2011) 1178–1199.
- [8] A. Abdel-Majeed, M. Braun, Low voltage system state estimation using smart meters, in: 47th International Universities' Power Engineering Conference (UPEC), 2012, pp. 1–6.
- [9] C. Fang, H. Xueshan, P. Zhiyuan, H. Li, State estimation model and algorithm including PMU, in: Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2008, pp. 1097–1102.
- [10] C. Bruno, C. Candia, L. Franchi, G. Giannuzzi, M. Pozzi, R. Zaottini, M. Zaramella, Possibility of enhancing classical weighted least squares State Estimation with linear PMU measurements, in: IEEE PowerTech, Bucharest, 2009, pp. 1–6.
- [11] P.S. Castro Vide, F.P. Maciel Barbosa, I.M. Ferreira, Combined use of SCADA and PMU measurements for power system state estimator performance enhancement, in: Third International Youth Conference on Energetics (IYCE), 2011, pp. 1–6.
- [12] A. Kumar, S. Chakrabarti, ANN-based hybrid state estimation and enhanced visualization of power systems, in: IEEE PES Innovative Smart Grid Technologies (ISGT), India, 2011, pp. 78–83.
- [13] C.N. Lu, J.H. Teng, W.H.E. Liu, Distribution system state estimation, *IEEE Trans. Power Syst.* 10 (1995) 229–240.
- [14] L. Ke, State estimation for power distribution system and measurement impacts, *IEEE Trans. Power Syst.* 11 (1996) 911–916.
- [15] A. Ranković, B.M. Maksimović, A.T. Sarić, A three-phase state estimation in active distribution networks, *Int. J. Electr. Power Energy Syst.* 54 (2014) 154–162.
- [16] I. Dzafic, M. Gilles, R.A. Jabr, B.C. Pal, S. Henselmeyer, Real time estimation of loads in radial and unsymmetrical three-phase distribution networks, *IEEE Trans. Power Syst.* 28 (2013) 4839–4848.
- [17] D.A. Haughton, G.T. Heydt, A linear state estimation formulation for smart distribution systems, *IEEE Trans. Power Syst.* 28 (2013) 1187–1195.
- [18] M.E. Baran, A.W. Kelley, State estimation for real-time monitoring of distribution systems, *IEEE Trans. Power Syst.* 9 (1994) 1601–1609.
- [19] M.E. Baran, A.W. Kelley, A branch-current-based state estimation method for distribution systems, *IEEE Trans. Power Syst.* 10 (1995) 483–491.
- [20] B. Golomb, T. Sejnowski, Sex recognition from faces using neural networks, in: A.F. Murray (Ed.), *Applications of Neural Networks*, Springer, US, 1995, pp. 71–92.
- [21] B.B. Thompson, R.J. Marks, M.A. El-Sharkawi, On the contractive nature of autoencoders: application to missing sensor restoration, in: International Joint Conference on Neural Networks, 2003, pp. 3011–3016.
- [22] V. Miranda, J. Krstulovic, H. Keko, C. Moreira, J. Pereira, Reconstructing missing data in state estimation with autoencoders, *IEEE Trans. Power Syst.* 27 (2012) 604–611.
- [23] J. Krstulovic, V. Miranda, A.J.A. Simoes Costa, J. Pereira, Towards an auto-associative topology state estimator, *IEEE Trans. Power Syst.* 28 (2013) 3311–3318.
- [24] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* 313 (2006) 504–507.
- [25] S. Narayanan, R.J. Marks II, J.L. Vian, J.J. Choi, M.A. El-Sharkawi, B.B. Thompson, Set constraint discovery: missing sensor data restoration using autoassociative regression machines, in: International Joint Conference on Neural Networks, 2002, pp. 2872–2877.
- [26] V. Miranda, N. Fonseca, EPSO—best-of-two-worlds meta-heuristic applied to power system problems, in: Congress on Evolutionary Computation, Honolulu, Hawaii, 2002, pp. 1080–1085.
- [27] V. Miranda, N. Fonseca, EPSO—evolutionary particle swarm optimization, a new algorithm with applications in power systems, in: IEEE/PES Transmission and Distribution Conference and Exhibition, Asia Pacific, 2002, pp. 745–750.
- [28] A.G. Madureira, J.A. Pecos Lopes, Coordinated voltage support in distribution networks with distributed generation and microgrids, *IET Renewable Power Gener.* 3 (2009) 439–454.
- [29] C. Gouveia, D. Rua, F. Ribeiro, C.L. Moreira, J.A.P. Lopes, INESC Porto experimental Smart Grid: enabling the deployment of EV and DER, in: IEEE PowerTech, Grenoble, 2013, pp. 1–6.
- [30] R.M. Ciric, A. Padilha Feltrin, L.F. Ochoa, Power flow in four-wire distribution networks—general approach, *IEEE Trans. Power Syst.* 18 (2003) 1283–1290.