2014 IEEE 8th International Power Engineering and Optimization Conference (PEOCO2014), Langkawi, The Jewel of Kedah, Malaysia. 24-25 March 2014

State Estimation in Distribution Smart Grids Using Autoencoders

[1]P. N. Pereira Barbeiro, [2]J. Krstulovic, [3]H. Teixeira, [4]J. Pereira, [5]F. J. Soares, [6]J. P. Iria

Instituto de Engenharia de Sistemas e Computadores, INESC TEC (formerly INESC Porto)
Porto, Portugal

[1]pnpb@inescporto.pt, [2]jopara@inescporto.pt, [3]hst@inescporto.pt, [4]jpereira@inescporto.pt, [5]fsoares@inescporto.pt, [6]jpiria@inescporto.pt

***Abstract*—This work proposes an innovative method based on autoencoders to perform state estimation in distribution grids, which has as main advantage the fact of being independent of the network parameters and topology. The method was tested in a real low voltage grid (incorporating smart grid features), under different scenarios of smart meter deployment. Simulations were performed in order to understand the necessary requirements for an accurate distribution grid state estimator and to evaluate the performance of a state estimator based on autoencoders.**

I. Introduction

State estimation is powerful technique that has been used by Transmission Systems Operators (TSO) since 1960s. It gives them a comprehensive and reliable view of the state of their networks in quasi-real time, allowing them to take the best decisions for maintaining power system operating properly. The state estimator uses the available measurements (voltage magnitudes, power injections and active and reactive power flows), network parameters and system topology and provides the best possible approximation for the state of the system, which means the determination of the system state variables: voltage magnitudes and phase angles in all buses.

In the last years, the growing deployment of Phasor Measurements Units (PMU) in power systems has changed the traditional paradigm of state estimation, usually based on asynchronous SCADA measurements with relatively slow refresh rates. Due to their characteristics, PMU can contribute to enhance significantly the performance and accuracy of the state estimation algorithms. Integrating PMU in the classical state estimation is a complex problem that has already been addressed in several published studies [1]-[4]. The advantages of PMU is that they are based on the application of precise time data provided by Global Positioning System (GPS) satellites, meaning that they can measure directly voltage phase angles and magnitudes with a very high accuracy. Yet, as PMUs are still very expensive, a state estimator based only in PMUs is not expectable in the near future.

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».

The most common technique for solving the state estimation problem is based on Weighted Least Squares (WLS) algorithm [5]. The WLS algorithm is a process that relies on the total knowledge of the grid (parameters and topology) and on the available measurements over a time-window and uses an iterative process to obtain the state estimation results. Therefore, the WLS algorithm usually consumes a significant amount of time (3-5 min), making it an impractical tool to be used for real-time visualization of the power system [6].

Even though state estimation is widely used for transmission grids, the first steps for its application at the distribution level were given just a few years ago. The reason behind this fact is that a critical event at the transmission level affects thousands of consumers, while in the distribution the amount of affected clients is substantially lower. Moreover, distribution network is much more complex and extensive than the transmission network and it would require an enormous investment in measurement equipment in order to guarantee the necessary system observability for performing accurate state estimation. Additionally, the lack of information about distribution grids, especially at Low Voltage (LV) level, also contributed to delay the application of state estimation in these systems.

The recent advent of the smart grid was the turning point regarding the possibility of including a state estimator in the distribution system. The smart grid concept is leveraging the development of new functionalities and the integration of new technologies in the power system, which can be used to foster the application of state estimation techniques in distribution grids (e.g. smart metering and associated communication platforms). Until now, very few works have exploited this topic in a holistic perspective. As an example, in [7], the authors tested the WLS technique considering a set of real time smart meters measurements. However, the technique used requires a topology processor (total knowledge of the grid parameters), which is inexistent in the majority of the cases. Furthermore, LV networks are usually unbalanced and classical WLS technique implies balanced loads.

This work addresses the problem differently, as it proposes an innovative methodology that is independent of the network parameters and topology. It uses the concept of artificial intelligence, through the application of autoencoders, to estimate the state of distribution grids. An autoencoder is a particular type of neural network frequently applied in areas related with pattern recognition and reconstruction of missing sensor signals [8], [9]. Recently, autoencoders were successfully applied in the power systems field for the detection of topological errors as well as for the generation of pseudo-measurements for classical state estimation methods [10], [11]. Nevertheless, until now, their use as "the core" of a state estimation algorithm was never implemented nor tested.

The methodology proposed in this work was tested in a real LV grid (incorporating smart grid features), under different scenarios of smart meters deployment (with the capability of transmitting data in quasi-real time). The main goals of the simulations performed were: a) understand the necessary requirements to have an accurate state estimator for distribution grids and b) evaluate the performance of a state estimator based on autoencoders.

II. THE CONCEPT OF AUTOENCODER

Auto-associative neural networks (AANN) or autoencoders are feedforward neural networks that are built to mirror the input space S in their output. The size of the output layer is the main difference between an autoencoder and a traditional neural network - in an autoencoder the size of its output layer is always the same as the size of its input layer. Therefore, an autoencoder is trained to display an output equal to its input. This is achieved through the projection of the input data onto a different space S' (in the middle layer) and then re-projecting it back to the original space S. With adequate training, an autoencoder learns the data set pattern and stores in its weights information about the training data manifold. The typical architecture of an autoencoder is a neural network with only one middle layer (Fig. 1). This simple architecture is frequently adopted because networks with more hidden layers have proved to be difficult to train [12], although allowing increasing accuracy.

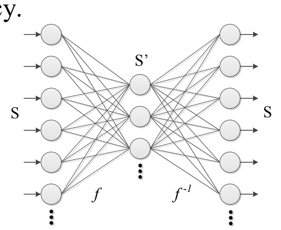

Fig. 1. Architecture of an autoencoder with a single hidden layer

There is no a priori indication of an adequate reduction rate (measured as the ratio between the number of neurons in the smallest middle layer and the number of neurons in the input/output layer) to be adopted. This decision on the reduction rate is dictated in present-day practice by trial and error and by characteristics of the problem. Autoencoders with thousands of inputs have been proposed for data or image compression, using the signals available in the middle layer, which maps the input to a reduced dimension space [8], [13], [14]. Once the autoencoder 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. Iteratively reintroducing the output value in the input will converge to a value that minimizes the input-output error. This approach is called Projection Onto Convex Sets (POCS) [15] and it uses alternating linear projections on the input/output space S to converge to the assumed missing value. A search may then be conducted by an optimization algorithm to discover the values that should be introduced in the missing components such that the input-output error becomes minimized. In the process denoted unconstrained search, the convergence is controlled by the error on the missing signals, whereas in the constrained search it is controlled by the error on all the outputs of the autoencoder. Any of these optimization procedures may be used, but the last one is the most suitable method to search a missing signal.

The application of autoencoders to power system problems is not very common. In [10] an EPSO is employed to recompose missing information in the SCADA of Energy/Distribution Management Systems (EMS/DMS) through the use of offline trained autoencoders. In [11], a model for breaker status identification and power system topology estimation is presented, but here, instead of a centralized/global estimation process, a mosaic of local estimators (based on a competitive auto-associative neural network principle) is used. More information about autoencoder applications can be found in [10].

III. METHODOLOGY

The methodology developed for solving the state estimation problem is based on the use of an autoencoder properly trained. A constrained search approach (as described in the previous section) is applied for finding the missing signals. Of course, in the context of state estimation, missing signals have to be necessarily the state variables – voltage values both in magnitude and phase (Fig. 2).

Within the constrained search approach, an Evolutionary Particle Swarm Optimization (EPSO) was chosen for reconstructing the missing state variables. The EPSO algorithm has been successfully applied already in several different areas of power systems. When it comes to its employment in autoencoders, some experiments were performed in [10] and the authors concluded that an optimization performed with EPSO (constrained search) was much more efficient than with others methods, such as POCS or unconstrained search.

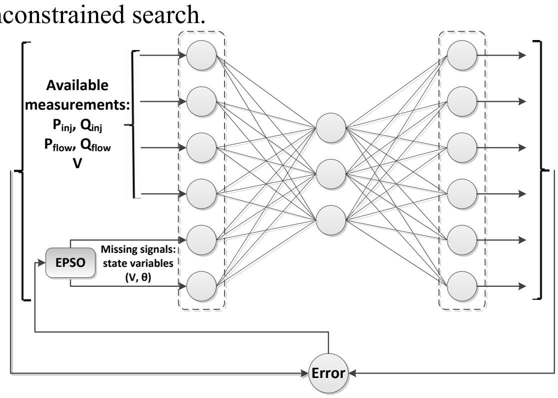

Fig. 2. Illustration of constrained search approach applied to the state estimation problem

Fig. 3 presents the flowchart of the main steps of the proposed methodology.

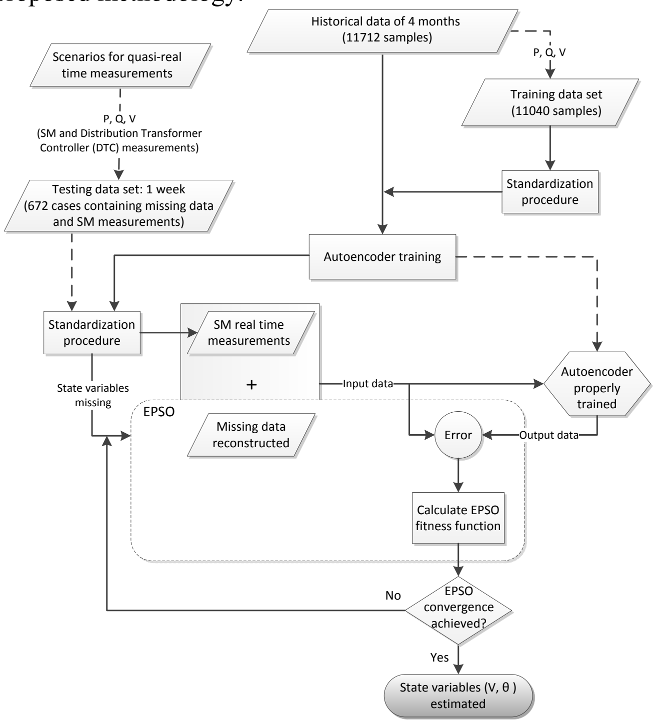

Fig. 3. Flowchart of the main steps of the proposed methodology

A. Historical Data

An effective state estimation through the use of autoencoders requires inevitably a large historical database, which needs to contain data about the variables that are passed to the autoencoder (missing signals and measurements). Additionally, the amount of data for each time instant/operating point should be available in enough number. This is crucial for a successful and effective training process since it is what enables the autoencoder to learn the necessary patterns/correlations between the electrical variables of a given network. There is no limit for the quantity of data in the historical database. However, it should be noted that the autoencoder learning process improves with the amount of data available. The only negative implication is that the training process will require more time.

A historical database with four months of real power flow results was created for testing the methodology. It required running several load flows in a typical Portuguese LV network in steps of 15 minutes for each day of the four months. Typical load patterns for summer months were considered. A full characterization of the study case, including the load and μG profiles adopted, is provided in section IV.

B. The Standardization Procedure

In order to have an efficient autoencoder, all the data must be standardized before being passed to it. As it can be seen Fig. 1, before the training and test process take place, a standardization procedure is run with the goal of pre-treating the input and output train data set. In this scale adjustment process, the range of the input and output values are transformed to a normalized interval of [-1,1]. This procedure allows a better adjustment of the input variables to the range of the activation function. Also, it allows the autoencoder to be less affected by the different ranges of the variables in the training data set. The method chosen here to perform the standardization was the "Min-Max method". This is the best standardization procedure when the minimum and maximum values of the data set are known. Looking to the historical database, the minimum and maximum values of the variables that compose the input vectors can be easily obtained.

C. Training Process

The autoencoder was trained with 11040 samples gathered from the database. These samples correspond to a total of 115 summer days. The input data set variables were settled in accordance with the scenarios defined in section IV. D. for the measurements that are transmitted in quasi-real -time. Regarding this aspect, it is important to state that whenever the quantity or type of measurement present in the input data set is changed, a new process of training must be performed.

In this study, the training was always performed with a number of neurons equal to the number of input variables, both at the input and output layer. In the hidden layer, the number of neurons were settled to be equal to 0.6 times the number of input variables (rounded to an integer value). The input variables are the sum of the measurements available plus the missing state variables.

A Resilient Back- Propagation algorithm was adopted for training the autoencoders properly. This algorithm 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 autoencoder the error is calculated, and then it is propagated back through the network while weights are adjusted in order to make the error smaller. The mean square error of the data set was chosen to measure the quality of the autoenconder during the training process. An interesting particularity of the training algorithm adopted is that instead of training combined data, the training data set is executed sequentially one input at a time, minimizing this way the mean square error for all training data set. At the same time, it provides a very efficient way of avoiding getting stuck in a local minimum. Some experimental training tests were carried out in order to select the most appropriate activation function for the hidden and output layer. After some analysis, a non-linear activation function, namely a symmetric sigmoid activation function, has proven to be the best choice for the activation function of both layers.

D. Testing the Autoencoder for Performing State Estimation

The testing phase consists of running the autoencoder, already trained, incorporating an optimization procedure for reconstructing the missing variables of the test set (see flowchart of Fig. 3). The fitness function of the optimization problem was defined to minimize the square error between the input and output of the autoencoder. Here, 672 samples of the historical database, corresponding to the last week of the last month, were tested. It should be referred that Gaussian noise was added to the test set measurements in accordance with the SM accuracy, as described in the end of section IV. B.

IV. Study Case

A. General Network Characterization

The methodology was tested in a typical Portuguese LV network where the MV/LV substation is equipped with one transformer of 100 kVA. A total of 57 consumers are present in the grid. Their contracted powers vary in a range between 3.45 kVA to 6.9 kVA for single phase consumers and 6.9 kVA to 13.8 for three-phase consumers. Since a significant amount of single-phase loads is present, load distribution among phases is not completely balanced. In fact, in some points of the grid, load imbalance is quite notorious. Even so, at the MV/LV substation load distribution is almost balanced.

In this study, the consumers' load was aggregated at the correspondent connection node and its distribution per phase was assumed to be completely balanced. Nevertheless, as this process is performed after using each individual consumer power value for a given time instant, the different consumers' load patterns are still reflected on the equivalent load. There are two main reasons for this simplification. First, by assuming balanced loads, single-phase load flows can be run instead of three-phase load flows. Second, the assumption made does not compromise in any way the quality of the state estimation results through the use of autoencoders.

Single-phase load flows was the process chosen in this study for generating historical and test data sets (both for training and validation phases). The single line diagram of the network used as test case is presented in Fig. 4.

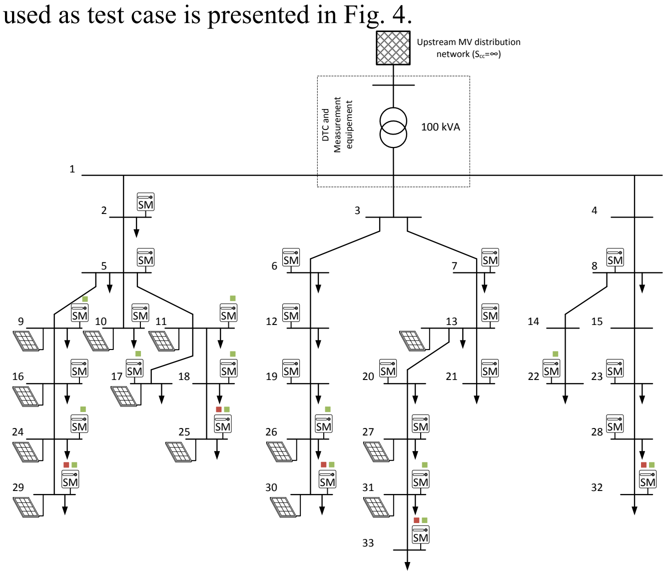

Fig. 4. Typical Portuguese LV network of 100 kVA used as test case

B. Description of the Smart Grid Features Added

In order to emulate a smart grid environment, some additional features and equipment were assumed to exist in this network. The MV/LV substation houses a Distribution Transformer Controller (DTC) as well as the associated measurement equipment which is capable of monitoring in quasi-real time the following variables: active and reactive power flows in the transformer and in all LV substation feeders and voltage (magnitude and angle) at the high and low voltage side of the transformer. 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 minutes (or even less, depending on the communication infrastructure).

μG units were added and distributed randomly through the network, totalizing 25% of the secondary substation transformer capacity (c.a. 25 kVA). The μG units were assumed to be photovoltaic panels and represent 50% of the contracted power of each consumer.

It was also considered that each customer has a SM to monitor his consumption and communicate it to the DTC for billing purposes. The customers that own a μG unit have an additional SM for measuring its power production. As it happens in some real smart grid test sites, not all SM are capable of transmitting data in quasi-real time due to communication infrastructure restrictions. Only some of the SM, which use, for instance, GPRS technology, have that capability. The selected SM that have this functionality, in each scenario analysed, are presented in section III. D. and it was assumed that their active power (P), reactive power (Q) and voltage (V) measurements are synchronised.

In terms of data accuracy, SM are usually categorized in classes, according to the confidence level specified by the manufacturers. After checking the technical specification of a large set of SM currently available in the market, some typical values were assumed. Voltage measurements were considered with ±1% accuracy and P and Q measurements with ±2%, all of them with a confidence level of 95%.

C. Load and μG Diagrams

Regarding load, the only data available for this grid was an average aggregated load diagram at the secondary substation level. In order to represent consumers with different behaviours, distinct load diagrams were generated from the aggregated diagram for each client. To this end, a Gaussian distribution with mean equal to the aggregated diagram and standard deviation of 8% was used. An example of the load profiles of 5 different customers, collected from the database are presented in Fig. 5. The load profiles are expressed in percentage of the peak load of the corresponding customer.

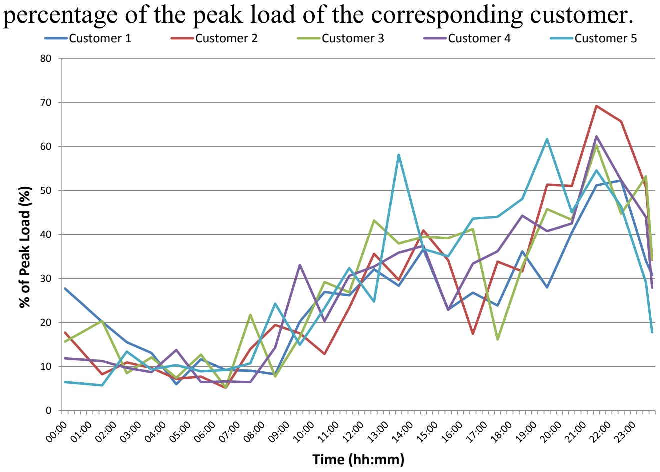

Fig. 5. Example of the load profiles of 5 different customers

Regarding the μG production diagrams, also aiming at representing different days (e.g. sunny, cloudy, rainy, etc.), 5 different real profiles obtained from a real meteorological station [16] were used and sorted randomly according to their probability of occurrence in a typical Portuguese summer.

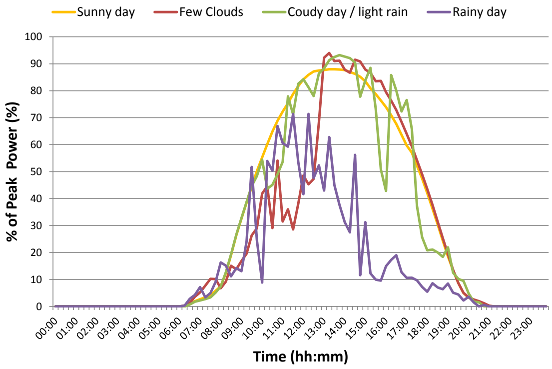

Fig. 6. μG production diagrams obtained from a real meteorological station

D. Scenarios for Quasi-real time measurements

Three scenarios were created to evaluate the performance of the autoencoders in the state estimation. In each scenario, the number of SM capable of transmitting data in quasi-real time was assumed to be different.

In scenario 1, no SM with capability of transmitting data in quasi-real time were considered – the only quasi-real time measurements are the P and Q power flows in the transformer and in all LV feeders and voltage (magnitude and angle).

Scenario 2 includes SM with capability of transmitting data in quasi-real time in the clients located in the farthest buses from the secondary substation.

In scenario 3, SM with capability of transmitting data in quasi-real time were continuously added to scenario 2, in the network buses with larger voltage magnitude estimated error, until the overall uncertainty remaining below a threshold of 2%. SM were added to the buses with larger voltage magnitude estimated error since it was assumed in this work that voltage magnitude is the most relevant output of a state estimator suitable for low voltage grids

An overall perspective of the SM location in each scenario is given in Fig. 4. The red dots show the location of SM in scenario 2 and the green dots show their location in scenario 3.

IV. SIMULATION RESULTS

The real and estimated voltage magnitudes, in the scenarios analysed, for the worst case of the test dataset (test sample with the biggest absolute error) are depicted in Fig. 2. Only buses without SM are shown, since in the others it was assumed that quasi-real time measurement are available (affected only by the Gaussian noise added). As it can be seen, estimated values are better when more quasi-real time measurements are considered. Looking to Fig. 4 and observing the buses distribution among the network, it is clear that the worst estimated values belong to buses that have both loads and micro-generation (*e.g.* buses nr. 17, 18, 25). This was an expected result due to the higher variability of the power injected these buses. Conversely, the developed state estimator performs better on buses belonging to the feeder located on the right side of the network (buses without any μG installed).

The variability of the voltage magnitude uncertainty for all the buses in scenario 3, considering the entire list of test samples simulated, is shown in Fig. 8. The green line represents the average value per bus, whereas the red lines represent a 2% uncertainty. Taking into account that scenario 3 was built in an attempt to reach a goal of 2% of uncertainty in the voltage magnitudes, the results depicted in Fig. 8 clearly demonstrate that the proposed methodology is able to successfully reach that target. Despite uncertainty being near to 2% in bus 16, the average value in the remaining buses is in the range +/- 0.5%. However, it should be noted that the 2% of uncertainty is only possible to reach when a total of 13 SM with the capability of transmitting data in quasi-real time are installed in the network. Even so, this value represents less than 50% of the total clients in the network. It is also important to stress that the amount of data present in the historical database may play an important role regarding the state estimation accuracy, since a larger historical database would lead to a more efficient training process.

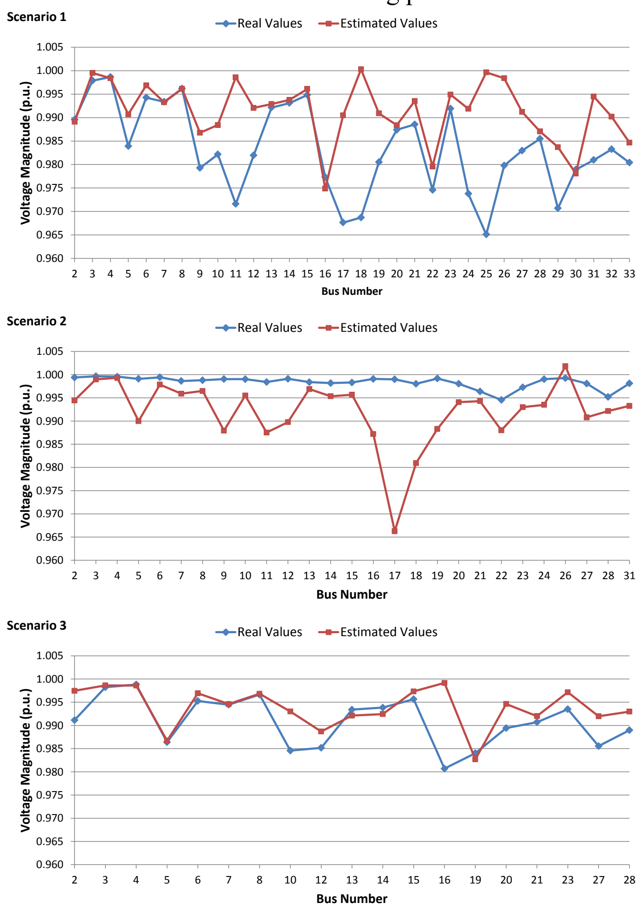

Fig. 7. Estimated and real values of voltage magnitude for the test sample with the highest absolute error in scenario 1, 2 and 3

Fig. 9 shows the maximum and average absolute error in the three scenarios analysed for both voltage magnitude and angle. As expected, a decrease in the error is verified in the scenario with more SM (see section IV. D.). In scenario 1, the maximum absolute error for voltage magnitude is ca. 0.035 p.u. (0.006 p.u. in average) and 0.56 degrees for the phase (0.08 degrees in average). When it comes to scenario 3, these values decrease to 0.018 p.u (0.002 p.u. in average) and 0.46 degrees (0.05 degrees in average).

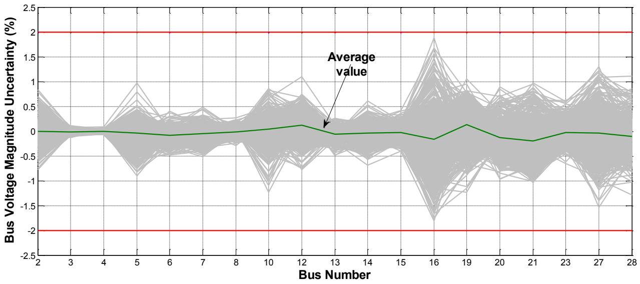

Fig. 8. Bus voltage magnitude uncertainty existent in scenario 3

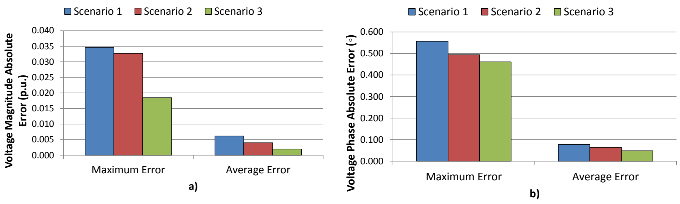

Fig. 9. Maximum and average absolute error of the a) voltage magnitude and b) voltage phase angle

The estimation error follows a Gaussian distribution both for voltage magnitude and voltage angle. An example of the error distribution is given in Fig. 10. In this specific case, the error follows a Gaussian distribution with an average value of 7.28×10^{-5} and standard deviation of 4.98×10^{-4}.

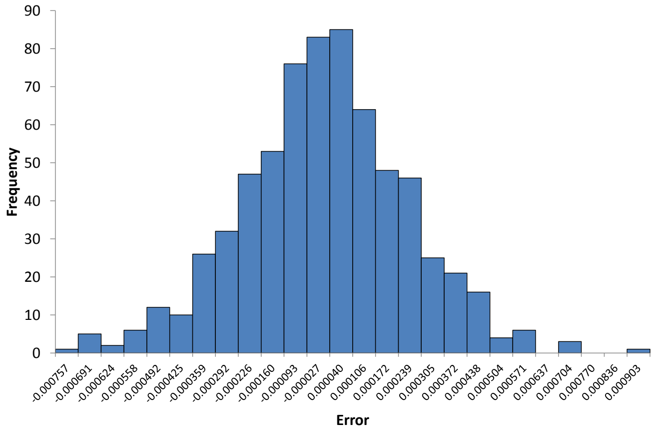

Fig. 10. Frequency of the voltage magnitude errors of one of the network buses in scenario 3

V. CONCLUSION

The smart grids paradigm envisages the existence of an advanced metering infrastructure capable of transmitting data in quasi-real time to grid operators. Since current smart meters are also capable of measuring the voltage values and active and reactive power flows, the referred infrastructure can also be used to transmit this data and serve as an enabler for the state estimation in distribution grids. Yet, one should bear in mind that an alternative to classic state estimation methods is necessary, since the distribution grids topology is not usually known, especially at the LV level. The methodology proposed in this paper, based on autoencoders, proved to be a very good alternative to the classic methods, since it allows overcoming this limitation. The results obtained from the simulations performed showed that when a large historical dataset exists and when enough SM are available, a state estimator based on the use of autoencoders may be very effective and accurate. Another strong argument in favour of the autoencoders is that they run in a very short period of time (only a few seconds), what makes them suitable for real time application.

VII. REFERENCES

[1] R.O. Burnett, M.M. Butts and T.W. Cease et al., Synchronized phasor measurements of a power system event, IEEE Trans. Power Syst. 3 (1994), pp. 1643–1650.

[2] T.S. Bi, X.H. Qin, Q.X. Yang, “A novel hybrid state estimator for including synchronized phasor measurements”, Electric Power Systems Research 78, pp. 1343–1352, Feb., 2008.

[3] S. Chakrabarti, E.Kyriakides, G. Valverde, V. Terzija, “State Estimation Including Synchronized Measurements”, PowerTech, Bucharest 2009

[4] P. Vide, M. Barbosa, I. Ferreira, “State estimation model including synchronized phasor Measurements”, UPEC 2011, Sept. 2011.

[5] A. Abur and A. Gómez Expósito, Power system state estimation: theory and implementation, New York, NY: Marcel Dekker, 2004.

[6] Amit Kumar and S. Chakrabarti, “ANN-based Hybrid State Estimation and Enhanced Visualization of Power Systems,” *ISGT India,* Dec. 2011

[7] Abdel-Majeed, A.; Braun, M., "Low voltage system state estimation using smart meters," 47th International Universities Power Engineering Conference (UPEC), September 2012.

[8] B. Golomb and T. Sejnowski, “Sex recognition from faces using neural networks,” in Applications of Neural Networks, A. Murray, Ed. Norwell, MA: Kluwer, 1995, pp. 71–92.

[9] B. B. Thompson, R. J. Marks, and M. A. El-Sharkawi, “On the contractive nature of autoencoders: Application to missing sensor restoration,” Int. Joint Conf. Neural Networks, Jul. 2003.

[10] V. Miranda, J. Krstulovic, H. Keko, C. Moreira, and J. Pereira, “Reconstructing missing data in state estimation with autoencoders,” IEEE Trans. Power Syst., vol. 27, no. 2, pp. 604–611, May 2012.

[11] Krstulovic, J. Miranda, V. Simoes Costa, A.J.A.; Pereira, J., "Towards an Auto-Associative Topology State Estimator," , IEEE Transactions on Power Systems, vol.28, no.3, pp.3311,3318, Aug. 2013

[12] G. Hinton, R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” Science, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[13] G. W. Cottrell, P. Munro, and D. Zipser, “Learning internal representations from gray-scale images: An example of extensional programming,” in Proc. 9th Annu. Conf. Cognitive Science Society, Seattle,WA, 1987.

[14] M. K. Fleming and G. W. Cottrell, “Categorization of faces using unsupervised feature extraction,” in Proc. IJCNN—Int. Joint Conf. Neural Networks, San Diego, CA, Jun. 17–21, 1990, vol. 2, pp. 65–70.

[15] S. Narayanan, R. J. Marks, II, L. Vian, II, I. I. Choi, M. A. El-Sharkawi, and B. B. Thompson, “Set constraint discovery: Missing sensor data restoration using auto-associative regression machines,” in Proc. Int. Joint Conf. Neural Networks, IEEE World Congr. Computational Intell., Honolulu, HI, USA, May 12–17, 2002, pp. 2872–2877.

[16] C. Gouveia, D. Rua, F. Ribeiro, C. L. Moreira, J. A. Peças Lopes, “INESC Porto Experimental SMART GRID: Enabling the Deployment of EV and DER”, Powertech 2013, June, 2013.