

Conceptual Analysis of Distribution System State Estimation of Low Voltage Networks

I. Táci¹, B. Sinkovics¹, I. Vokony¹ and B. Hartmann¹

¹ Department of Electric Power Engineering
Budapest University of Technology and Economics
Egry József str. 18., Budapest, 1111, Hungary
Phone/Fax number: +0036 5166473, e-mail: taczi.istvan@vet.bme.hu

Abstract. Distribution System State Estimation has great potential in the integration process of renewable energies, energy storage and active participation of the customers. The international state of the art pilot projects showed that there are many use cases, such as indication of operational limit violations and enhancing strategic development and asset management by useful in-depth information about the electrical behaviour of the system. However, the concept of state estimation at the distribution level is still a developing research area due to the structural differences of the network structure and parameters from the transmission applications which is also discussed in detail. Based on the value proposition identified through the literature review, simulation studies are carried out on distribution test system with different scenarios based on renewable penetration levels and measurement data availability. The main contribution of this research is a conceptual overview of the possible application of distribution system state estimation and validation of the operation on a case study, which could serve as a foundation for testing the proposed viable use-cases.

Keywords. Distribution System State Estimation, State Estimation, Renewable Integration, Active Distribution Systems

1. Introduction

Nowadays, distribution system operators (DSOs) face challenges in the integration of numerous new technologies and processes. Distributed generation (DG) brings structural changes in the generation mix by connecting to the medium voltage (MV) and low voltage (LV) networks. The active participation of customers through demand side management (DSM) processes – either controlled in active systems or incentivized by tariff schemes – with the presence of new, significant electric power consumers (e-mobility chargers and heat pumps) basically change the load behaviour. Distributed energy storage systems also appear due to the more competitive nature of the renewable support schemes. These trends point toward an active distribution system (DS), therefore more sophisticated operation and control approaches are inevitable.

One of the most promising directions in developments is the application of state estimation (SE) which is by definition “a data processing algorithm for converting

redundant meter readings and other available information into an estimate of the state of an electric power system” [1]. Distribution system state estimation (DSSE) has great potential to increase observability and controllability at DS level, while still have many open questions [2] and characteristic differences compared to the widely used high voltage (HV) transmission system state estimation (TSSE), as will be discussed in in this paper.

Due to the limited measurement coverage of the DS, early studies were published in the 90’s. These papers have already used pseudo measurements to satisfy observability criteria [3][4].

If there are multiple sources of erroneous measurements or the redundancy is insufficient, the state estimation usually fails to converge. For that reason, the input set is synthetically expanded by introducing pseudo-measurements, which are created according to the network’s historical data by using some of the probabilistic and neural networks-based methodologies [5][6].

Then observability analysis aims to determine whether a state estimation solution for the specific system exists, the state variables can be calculated from the available measurement set. Moreover, this procedure identifies unobservable parts of the network which are a consequence of bad or missing data [1].

There are several different observability approaches, which can be divided into probabilistic [7][8] and machine learning based approaches [9]. All these methods rely on changing the bad data into pseudo-measurements generated by the different techniques.

If the network is adequately defined, the state estimation solver serves to find an optimal solution for the system states according to the network model constraints and given measurements. The optimal solution essentially comprises of a vector of complex bus voltages, from which estimates are calculated for other network values. The linear optimization solver usually includes a linear optimization problem (e.g. Weighted Least Squares) and an iterative objective function minimizer [10][11].

This research summarizes state of the art DSSE practical approaches from a conceptual point of view and validates the application for possible use cases through simulation

studies with different scenarios. The paper is organized as follows: section 2. briefs an overview of practical application examples, while section 3. introduces the DSSE methodology aspects and the modelling considerations of this study. Section 4. describes the simulation scenarios and results, and Section 5. summarizes the main conclusions.

2. International Examples

Research institutes and DSOs around the world already work on pilot applications of DSSE with promising results. Firstly, a French project called *evolvDSO* is introduced, where two novel algorithms were developed to apply DSSE in MV and LV networks for Enedis DSO. [13]

This tool is useful for both the transmission system operator (TSO) and DSO, as the calculation of flexibility ranges at primary substation level, while also providing voltage profiles for the considered LV network by using artificial neural networks.

The MV level algorithm called interval constrained power flow improves the system security by estimating the available flexibility for the following day. The algorithm considers the planned switching state (topology), the local consumption and generation forecasts, grid and asset constraints, flexibility resources (on-load tap changing transformers, static reactive power control devices, firm and non-firm connection contracts and market based bids from MV generators and battery storage systems) available at DSO level and the associated costs.

The tests showed positive results as the tool fulfilled the expectations, provided the range and place of flexibility in a reasonable time. The LV tool evaluates the voltage profile of an MV/LV supply area (around 1250 LV customers) using historical measurements and assumptions for forecasts.

The algorithm is a neural network-based approach and uses the phase connection data for customers, smart meter voltage and power measurements from an 18-month period. The analysis also considers the optimal number and placement of real-time measurement data usage. The calculations seem to have acceptable level of error, however the photovoltaic (PV) generation was neglected. The main lessons learned are that synchronism of the measurements is critical, while only 1 real-time measurement per feeder per phase is sufficient for the majority of improvement in accuracy, and the sensitivity for placement is low. As a possible future work, the authors propose to use PV data and enhance synchronism of measurements which is clearly critical.

In a Scottish pilot project from Orkney Isles a new error estimation technique was proposed. [14] The main motivation behind the project was the growing share of DG (mainly wind – which has a capacity factor of 40% in the area and also wave and tidal energy from the local European Marine Energy Centre) in the area.

The focus was to find an appropriate mix of measurement and pseudo-measurement with acceptable error rates. The constraints were twofold: on one hand it was usually thermal limits before the DG level increase, but on the other hand voltage variations also became problematic since then. The network model had 70 buses and 76 branches, from which 15 busbar voltages and 30 current

flows were visible and with a small number of pseudo-measurements, 100% visibility was reached at the MV level. As a result, the DSO have been connected further 20 MW of DG to a network that said to have reached the hosting capacity limits.

A novel decoupled method for the power flow equations is used in DSSE algorithms by using the characteristics of a usual DS, namely the high R/X ratio. [15] Therefore, the system is the opposite of TS – the relation between active power and the angle (P- θ) and the reactive power and voltage (Q-V) due to the reactive lines can be transformed into a P-V and Q- θ in this case.

The authors tested the method on a Slovenian unbalanced network example and compared the classical WLS to the simplified (decoupled) WLS. The proposed method has significantly reduced computation time (around -46%) while the errors remained the same.

An intriguing example of LV DSSE based on smart meter data is the SmartSCADA project in Kaiserslautern, Germany [16]. The assumption was that the smart meters at the residential customers could offer enough visibility for the DSSE. The LV network consisted of 120 loads and 24 PV generators, and 110 smart meters were installed with the possibility to send voltage, current magnitudes as well as active and reactive powers through power line communication. A linear algorithm was developed to reduce to computation burden caused by the need for a three-phase estimation to calculate each phase correctly, with the assumption that the relative angles are constant and symmetrical, only magnitude differences are possible.

Plausibility checks are also included from operational constraints such as maximum line current of the fuses. The field test results showed promising results, the voltage magnitude estimation accuracy was within a range of +/- 1.4 V, which is enough precision to localize possible voltage limit violations on the grid. The increased measurement redundancy due to the smart meter rollout led to a conclusion that the erroneous measurements can be detected in the voltage, and most cases in the current as well (only low loading condition is problematic for the proposed algorithm).

Another German example is from Aachen, where an in-field test of meter placement and application of different DSSE algorithms were carried out at MV level [17]. The main scope of the project was to handle the observability problem by the lack of reliable measurement points. Pseudo-measurements were generated from different data, such as load profile databases, weather data, distributed generation nominal values and historical values (e.g. highest load), while the error was calculated by Monte Carlo simulation approach.

The scope was to reach the limits defined by EN50160 standard (99% quantile for voltage, 95% quantile for current estimation errors). The meter placement was carried out by a genetic algorithm with the objective function to minimize the investment costs.

The tests were carried out at an MV network with 14 MV/LV substations, where 13 of them had measurements from three-phase currents, voltages, active and reactive power. Different communication technologies were used (Modbus RTU RS485, power line communication and radio) to reach a central protocol converter, then IEC

61850 will be the final solution at the end of the pilot. The error from the historical highest load pseudo-measurements was high (around 29%), while the enhanced pseudo-measurements (considering the effects of distributed generation) brought it down to 5% average. As a summary, the above-mentioned projects offer viable use-cases with pilot application results, such as flexibility range calculation and voltage limit violation indication [13], distributed generation hosting capacity increase [14]. Different approaches are present to generate pseudo-measurements, such as using historical data, predicted load and generation profiles and weather data. The impedance characteristics [15], availability of useful data from smart meters around the grid [16] and the effects of distributed generation [17] are all challenges in the practical application.

3. Methodology

DSSE differs highly from TSSE, therefore the appropriate methods are also distinct. This section summarizes the main challenges, compare the possible algorithms and propose a feasible modelling approach to analyse use cases.

A. Theoretical challenges of DSSE

An obvious speciality of DSSE is the observability problem: unlike the transmission level, where usually almost all of the bus voltages are measured, DS has much lower observability level – it is a common practice that only the HV/MV substations provide reliable data, while MV and LV network parts are not visible real time for the operators. Communication systems also constrain DSSE due to bandwidth, transfer capacity and endpoint handling limits and cybersecurity issues [18]. So, the first challenge is to deal with the limited availability of measurement points due to the limited amount of data and increased uncertainty. Different solution approaches, such as the use of pseudo-measurements (e.g. pseudo power injections at feeder buses with Gaussian distributions with their means at half the transformer rating etc.), smart meter data (which has considerable delays) and virtual measurements (e.g. zero power flows on open switches etc.) was introduced in the literature to handle this issue [3].

The topological attributes of DS are also important to consider, such as [1][2][16][19]:

- the high R/X value (resistance to reactance ratio) which means that conventional direct current (DC) (where resistances are neglected in the calculations to increase computing speed) SE algorithms are unusable
- higher complexity due to the unbalances in the system,
- much larger network with model uncertainties,
- low amount of data about the exact connecting elements, such as distributed generation and individual loads.

These challenges need adequate, robust estimation algorithms which are suitable to handle the diverse sources of data beside the high uncertainty and complexity, therefore DSSE is still an intriguing research field.

B. Comparison of SE algorithms

Based on timing, evolution and purpose of the actual estimation, SE methods can be classified into different categories. Table I. summarizes some of the most common mathematical methods with the main advantages and disadvantages.

Table I. – Summarization of the advantages and disadvantages of different SE methods [20][21][22]

Method	Pros and Cons
Weighted Least Square (WLS)	(+) Fast, simple, widely-used, (-) Sensitive to bad data
Least Median of Squares (LMS)	(+) Robust against bad data and leverage points, (-) High computational cost, high redundancy requirements
Least Trimmed Squares (LTS)	(+) Robust against bad data, (-) High computational cost and memory requirement
Least Absolute Value (LAV)	(+) Robust against bad data, small sensitivity to line impedance uncertainty, (-) High computational cost, sensitivity to leverage points and measurement uncertainty
Generalized Maximum Likelihood (GM)	(+) Robust against bad data, (-) Parameter selection sensitivity

Softwares used for power system analysis (DigSilent Power Factory, NEPLAN, PSS/E etc.) have the possibility to use different calculation methods of SE. While WLS has the disadvantage in robustness, it is still a fast and widely-used approach and going to be used in this study too. In the case of a DSSE solution is tailored to a practical application, other algorithms should be considered.

C. Modelling approach

The scope of this research is to validate the concept of DSSE as a useful tool to enhance the value propositions which are considered viable from the pilot projects through simulation studies. The paper investigates the practical challenge of SE simulation studies the lack of grid/load measurements. The model represents a usual MV/LV transformer area of Hungary, derived from the local DSO's anonymous and aggregated dataset as follows:

- topology, impedance parameters and ratings,
- active and reactive power flows on the terminal,
- measurement data (voltage, current), characteristics and locations at the end of the circuits and on the terminal.

The first step is to create a static DSSE which is able to create reliable models for symmetric alternating current (AC) load flow calculations. Based on the literature examples given in section 2, the inclusion of DG and the availability of measurements for appropriate observability and possible simplifications to reduce computation time is also considered. This means that the main contribution of this research is the application on an

exact LV network, assessment of field measurements and usual circumstances providing useful conclusions for the DSO.

The topology was built based on the database (transformer parameters, line length and impedances, load and generation connection points, rated powers). Generators are static ones with unity power factor, loads are modelled as constant power (consumption does not depend on node voltage) using the regular impedance-current-power (ZIP) static models. LV protection devices (fuses) are also included in the model with ratings. Built-in WLS DSSE algorithm was used in DigSilent PowerFactory 19.

The main limitations of this case study are the following:

- only some special system states are analysed, the effect of time sweep calculation is not discussed, nor any dynamics
- balanced network conditions mean that asymmetry is neglected.

However, future work could include the analysis of stochastic load and generation profiles. Based on the measurements, modelling of unbalanced conditions is also possible.

4. Simulation studies

In this section, the simulation scenarios and the results are introduced in detail. The test system is the supply area of an MV/LV transformer (depicted in Fig. 1.). Table II. summarizes the main topological parameters of the location. The load and generation profiles are assigned from a predefined database of the DSO's define. For the simulations, DigSilent Powerfactory 19 software was used. This location perfectly represents a usual developing small town in Hungary. Distributed generation is already present at the location, 12 PV systems are operating in the MV/LV area.

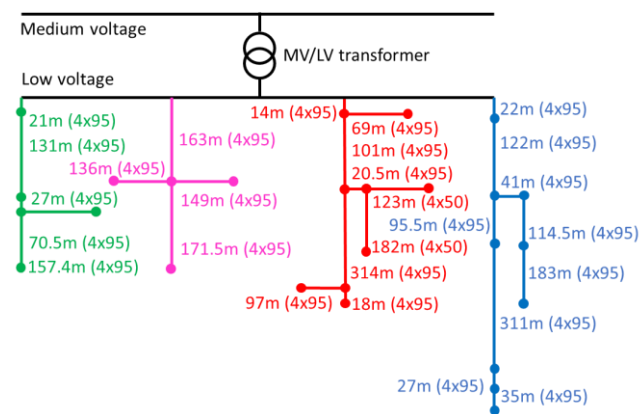


Fig.1. Topology of the considered test system

Table II. – Main topological parameters

Circuit	Total length [m]	Number of customers [-]	Number of PV systems [-]	Cumulative PV power (kW _p)
1	407	32	1	2
2	620	51	7	37.2
3	940	54	0	0
4	952	71	4	26.3

There are available measurements from the LV terminal of the transformer (voltage, current magnitude, active and reactive power) and the endpoints of circuit 1,2 and 4 (voltage magnitude). 12 PV systems are operating at the site currently, 7 of them on the B2 circuit, therefore the effects of the DG are mostly visible at that network part. B4 is the longest with the most customers, so the voltage drop is the largest at that part. Fig 2. represents the measured voltage values. The time-step of the measurement was 10 min, and the values were averaged (3 samples for the same day from the 3-week measurement period). This dataset was used as an input for the DSSE algorithm of PowerFactory.

Since the data about phase connections for both loads and generations is unavailable, balanced conditions were assumed. However, using further measurements and a location check it would be possible to extend the model to unbalanced studies. From the load database a high load condition was composed, the power factor of the loads is constant 0.95.

Based on the measurements that could be considered as daytime at the location. Therefore, the PV systems could be modelled with the kW_p infeed expecting optimal operating conditions to maximum effect of the DG. A unity power factor was considered, this is the defined values in the grid codes and regulations for household-sized DG (under 50 kVA) in Hungary. There are 3 different overhead lines, 2 4×95mm² diameter with the R/X ratio of 4.38 and 0.915, and the 4×50mm² diameter with 1.6 respectively. These R/X ratios mean that the decoupling of P-Q flows cannot be modelled by impedance simplifications here in the case of LV overhead lines.

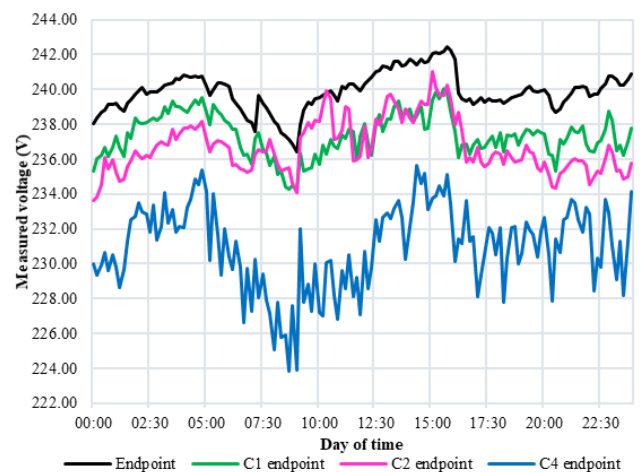
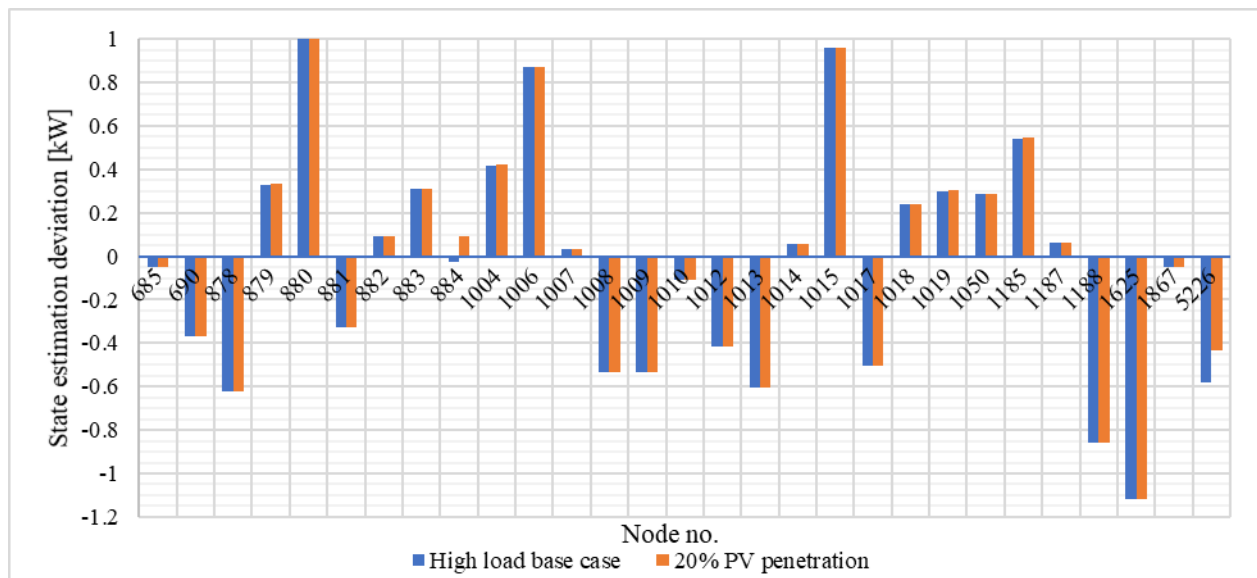


Fig.2. Averaged measured values from the location

There are two study cases with respect to the nature of measurement data which were considered:

- base case behaviour of the DS,
- increased number of PV generators

The scenarios are assigned with pseudo-measurement techniques as well to create the DSSE. The simulations aim check the distribution system state estimation capabilities of the Newton-Raphson calculation to solve the minimization problem by using Lagrange multipliers.



Node no.	685	690	878	879	880	881	882	883	884	1004	1006	1007	1008	1009	1010
High load	3.80%	20.22%	23.52%	-18.64%	-35.87%	26.11%	-9.98%	-20.16%	0.78%	-13.28%	-29.16%	-3.82%	30.18%	30.08%	3.82%
20% PV	3.73%	20.11%	23.48%	-18.70%	-35.27%	26.03%	-10.09%	-20.22%	-2.52%	-13.34%	-29.19%	-4.07%	30.07%	29.97%	3.79%
Node no.	1012	1013	1014	1015	1017	1018	1019	1050	1185	1187	1188	1625	1867	5226	
High load	20.16%	23.54%	-18.67%	-35.72%	26.08%	-9.69%	-20.08%	-13.30%	-29.14%	-3.86%	30.18%	30.06%	3.84%	31.29%	
20% PV	20.11%	23.47%	-19.00%	-35.76%	26.03%	-9.73%	-20.15%	-13.34%	-29.25%	-3.92%	30.11%	30.03%	3.76%	23.44%	

Fig.3. Load flow and State estimation difference on each node
Table III. – Relative state estimation error on each node

Each of the two case studies includes three already defined major SE functionalities, namely Plausibility check, Observability Analysis and Bad Data Detection.

The State Estimator function of PowerFactory provides consistent load flow results for an entire power system, based on real time measurements, manually entered data and the physical constraints of network model topology.

A sufficient network parametrization requires both metered node data and power flows of energy producer and consumer entities. In this paper a practically common incomplete metered dataset was used. The system operator provided metered power, voltage and current data from the endpoints of circuits and the MV/LV transformer terminal. For this reason, a household load prediction procedure was used to estimate household pseudo measurement values which represents the actual network state.

A normalized load profile database was used, the loads were paired with each profile using a random shuffle algorithm implemented in MATLAB. Then the similar point of time load was used as a pseudo measurement value for each load. The DG entities were parametrized the same way using normalized solar power generation profiles.

A. High load case study

The first case scenario introduces a base simulation of the described LV network without photovoltaic power generation and high energy demand. Due to the large number of implemented loads only circuit 1 is detailed in this paper. The state estimation results represent the grid condition on a weekday at 15:30. Fig. 3 depicts the load P estimation differences (blue) using load flow and state estimation, practically this graph reveals the accuracy of

the load profile generation too. Table III. contains information to assess the relative deviation of state estimation regarding load flow results.

B. Increased number of PV generators

Due to the integration of intermittent distributed generation and bidirectional power flows distribution system states have become more dynamic. For this reason, the second scenario represents a 20% PV penetrated grid case, namely 20% of nominal load is covered by installed PV capacities on the grid. The results are depicted in Fig. 3. using orange bar color.

As it is seen both cases show a relatively good approximately 18% average load estimation accuracy without using any complex (e.g neural networks) load forecast methods. The simulation results prove the preconception that the introduced pseudo measurement prediction method generates reliable dataset without outlier values.

5. Conclusions and remarks

This research evaluates the concept of DSSE by summarizing the experiences from current pilot projects and comparing the theoretical approaches from the current literature. Due to the increasing need for controllability and visibility of distribution systems, DSSE applications have many use cases, such as:

- integrating weather data for forecasting and smart meters,
- indicating operational limit violations, such as unacceptable voltage values, overloading,

- offering in-depth, standardized data structures for planning and development studies,
- providing input for smart asset management calculations.

The main purpose of this research is to validate use the DSSE process on a test system with appropriate modelling assumptions to support further studies and provide an application environment for DSO use-cases. Other contributions are further discussions on recently published pilot projects from the viewpoint of comparability and replicability, while also adding valuable analysis on the current directions and open questions regarding DSSE. The results of the simulation studies offer applicable results on the effects of visibility (measurement availability constraints) and provide a fundamental conceptual analysis for further in-depth use-case analysis.

Acknowledgement

This research was supported by the Hungarian Academy of Sciences and E.ON Hungary through the “Célzott Lendület” Program application of the MTA-BME FASTER Research Group.

References

- [1] A. Primadianto and C-N. Lu, “A Review on Distribution System State Estimation”, *IEEE Transactions on Power Systems* (2019), Vol. 32, pp. 3758-126.
- [2] G. Wang, G. B. Giannakis, J. Chen and J. Sun, „Distribution system state estimation: an overview of recent developments”, *Frontiers of Information Technology & Electronic Engineering* (2019), Springer, Vol. 20., no. 1., pp. 4-17.
- [3] M. E. Baran and A. W. Kelley, “State estimation for real time monitoring of distribution system,” *IEEE Trans. Power Syst.* (1994), vol. 9, no. 3, pp. 1601–1609
- [4] I. Roytelman and S. M. Shahidehpour, “State estimation for electric power distribution systems in quasy real time conditions,” *IEEE Trans. Power Del.* (1993), vol. 8, no. 4, pp. 2009–2015
- [5] J. Wu, Y. He and N. Jenkins, “A robust state estimator for medium voltage distribution networks” *IEEE Trans. Power Syst.* (2013), vol. 28, no. 2, vol. 28, no. 2, pp. 1008-1016
- [6] L. Mili, M. G. Cheniaie, N. S. Vichare and P. J. Rouss, “Robust state estimation based on projection statistics [of power systems]” *IEEE Transactions on Power Systems* (1996), vol. 11, no. 2, pp. 1118-1127
- [7] R. Singh, B. Pal and A. J. Rabih, “Statistical Representation of Distribution System Loads Using Gaussian Mixture Model” *IEEE Transactions on Power Systems* (2010), vol. 25, no. 1, pp. 29-37
- [8] Duy Thanh Nguyen, “Modeling load uncertainty in distribution network monitoring” *IEEE Transactions on Power Systems* (2015), vol. 30, no. 5, pp. 2321-2328
- [9] E. Manitsas, R. Singh, P. C. Bikash and G. Strbac, "Distribution System State Estimation Using an Artificial Neural Network Approach for Pseudo Measurement Modeling" *IEEE Transactions on Power Systems* (2012), vol. 27, no. 4, pp. 1888-1896
- [10] A. Minot, Y. M. Lu and N. Li, “A Distributed Gauss-Newton Method for Power System State Estimation” *IEEE Transactions on Power Systems* (2016), vol. 31, no. 5, pp. 3804-3815
- [11] T. Ratnarajah, Z.-Q. Luo and K. Wong, “Semidefinite programming solutions to robust state estimation problem with model uncertainties”, 1998
- [12] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan and F. Bu, “A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems” *IEEE Transactions on Smart Grid* (2019), vol. 10, no. 2, pp. 2312 - 2322
- [13] M. V. Sebastian, M. Caujolle, B. G. Maraver, J. Pereira, J. Sumaili, P. Barbeiro, J. Silva and R. Bessa, “LV state estimation and TSO-DSO cooperation tools: results of the French field tests of the evolvDSO project”, *Open Access Proceedings of the 24th International Conference & Exhibition on Electricity Distribution (CIRED)* (2017), Glasgow, pp. 1883-1887.
- [14] E. Davidson, A. Keane, R. Currie, N. Mcneill, D. Macleman and M. Lee, “Requirements-driven Distribution State Estimation”, *Open Access Proceedings of the 22nd International Conference on Electricity Distribution (CIRED)*, Stockholm, Paper no. 1100.
- [15] M. Antoncic and B. Blazic, “LV Network State Estimation Using Decoupled Load-flow Algorithm”, *Open Access Proceedings of CIRED Workshop* (2018), Ljubljana, Paper no. 287.
- [16] D. Waersch, R. Brandalik, W. H. Wellssow, J. Jordan, R. Bischler, N. Schneider, „Field Test of a Linear Three-phase Low Voltage State Estimation System Based on Smart Meter Data”, *24th International Conference on Electricity Distribution*, Glasgow, 2017.
- [17] D. Echternacht, M. Franken, A. Moser, U. Geulen, P. Zimmer, W. Feilhauer, W. Schrieder, „Smart Area Aachen – In Field Test of Meter Placement and State Estimation Algorithms for Distribution Grids, *IEEE PES Innovative Smart Grid Technologies Latin America*, 2015.
- [18] Y. Nie, C. Y. Chung, and N. Z. Xu, “System state estimation considering EV penetration with unknown behavior using quasi-Newton method,” *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4605–4615, Nov. 2016.
- [19] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaahedi, “Distribution system state estimation based on non-synchronized smart meters,” *IEEE Trans. Smart Grid* (2015), vol. 6, no. 6, pp. 2919–2928.
- [20] U. Kuhar, M. Pantos, G. Kosec, and A. Svigelj, “The impact of model and measurement uncertainties on a state estimation in three-phase distribution networks,” *IEEE Transactions on Smart Grid* (2019), Vol 10., No. 3., pp. 3301-3310.
- [21] M. Gol and A. Abur, “LAV based robust state estimation for systems measured by PMUs,” *IEEE Transactions on Smart Grid* (2014), vol. 5, no. 4, pp. 1808–1814.
- [22] J. Zhao, G. Zhang, M. L. Scala, and Z. Wang, “Enhanced robustness of state estimator to bad data processing through multi-innovation analysis,” *IEEE Transactions on Industrial Informatics* (2017), vol. 13, no. 4, pp. 1610–1619.