State Estimation for Medium and Low Voltage Distribution Grids Based on Near Real-time Grid Measurements and Delayed Smart Meters Data

Mohammad RAYATI, Thomas PIDANCIER, Mauro CARPITA, Mokhtar BOZORG
UNIVERSITY OF APPLIED SCIENCES OF WESTERN SWITZERLAND (HES-SO)
Route Cheseaux 1, Case Postale 521, 1401 Yverdon-les-Bains, Switzerland
e-mail : mokhtar.bozorg@heig-vd.ch
URL: http://iese.heig-vd.ch/

Acknowledgements

This research project is financially supported by the Swiss Innovation Agency Innosuisse and is part of the Swiss competence Center for Energy Research SCCER FURIES. The authors would also like to acknowledge the technical support of “Services Industriels de Genève - SIG”.

Keywords

«Estimation Technique», «Distribution of electrical energy», «Measurement», «Smart grids».

Abstract

In this paper, first, a distribution system state estimation (DSSE) algorithm based on Distflow model is presented. The model is designed for radial low voltage (LV) distribution grids considering traversal components such as the cable capacitances. Next, to obtain the pseudo-measurements required for the DSSE algorithm, we developed and compared three intraday nodal load forecast (INLF) methods that take; i) near real-time stream of data from the grid measurement devices (MDs), i.e., voltage and current magnitudes at limited number of lines/nodes, and ii) the batch data set from smart meters (SMs) available for previous days. The accuracy associated with the INLF methods is quantified in terms of statistical characteristics such as average of symmetric mean absolute percentage error (sMAPE) over all nodes of the distribution grid. Afterwards, the impact of this accuracy on the DSSE results is investigated using real data available for a distribution grid in city of Geneva, Switzerland.

1 Introduction

The recent increase of uncontrolled renewable production connected to the distribution network raised the need for better network monitoring and control solutions. Regulators encourage the distribution system operators (DSOs) to improve the network observability by installing low voltage (LV) grid measurement devices (MDs) as well as smart-meters (SMs) on the end-user clients’ side. For instance, in Switzerland at least 80% of the DSOs’ clients should be equipped with SMs by the end of 2027 [1]. The medium voltage (MV) and LV distribution networks are very large in terms of the number of lines/nodes. Therefore, the DSOs are looking for monitoring and supervision systems that require a limited number of affordable MDs [2]. In addition, due to the privacy issues, the measurements coming from MDs installed on the grid and those from the SMs at customer level are not available with the same sampling time and recuperation delay. In particular, in Switzerland the SMs data (i.e., the active/ reactive power consumptions of end-user clients at every 15-minutes interval) are not available in real-time and can only be recuperated once every 24 hours. Hence, the distribution system state estimation (DSSE) algorithms and intraday nodal load forecast (INLF) methods are required to improve the quality of network supervision with a limited number of MDs.

A number of DSSE algorithms are proposed in the literature. These algorithms differ in terms of required measurement data, network model, load flow equations, and estimation technique. A comprehensive review of the state of the art is presented in [3]. Most of the algorithms use the voltage angles (i.e., phasor measurement data), which are not always available with high level of accuracy. Indeed, the distribution grid lines are short; hence, the difference between voltage angles of two neighboring nodes can be as small as the MDs accuracy level. In [4], a DSSE algorithm is presented relying on the solution of Distflow equations [5] to estimate the magnitudes of voltages and line currents. This approach is appropriate for LV distribution grids since it does not require any voltage angle measurements or estimations. However, the shunt components (e.g., cable capacitance) of the lines are not taken into account. In [6], the authors have enhanced the DSSE algorithm proposed in [4] considering the shunt components of the network and using a PI model of the lines.

A DSSE algorithm estimates the state of the grid using: i) the voltage magnitudes and active/reactive power flow measurements (e.g., 1-hour average values) coming from the MDs installed at specified nodes/lines; and ii) the load pseudo-measurements obtained from the intraday forecast of active/reactive power injection/consumption at every node (i.e., INLF).

For benchmarking purposes, we developed the following three INLF methods to quantify the impacts of pseudo-measurements error and eventually the error of the DSSE:

- Seasonal auto-regressive integrated moving average (SARIMA) method using SMs data,
- Random forests method using SMs data,
- Random forests method using SMs and MDs data.

The SMs data includes historic time-series of aggregated active/reactive power consumption profiles at each node of the grid from the previous days (excluding intraday profiles that are not available due to privacy and legal issues). The MDs data includes historic and intraday time-series of voltage magnitude, active/reactive power measurements at specified nodes/lines where MDs are installed.

The rest of this paper is organized as follows: The enhanced DSSE algorithm is briefly presented in Section 2. The INLF forecast methods are discussed in Section 3. Afterwards, a case study is developed and described in Section 4 in order to demonstrate the effectiveness of the DESS algorithm and compare different INLF methods. The numerical results are presented and discussed in Section 5. Finally, summary and conclusions are given in Section 6.

2 Distribution System State Estimation

The DSSE algorithm used in this paper is proposed by the authors in [6]. Indeed, it is an improvement of the DSEE algorithm based on Distflow equation that is originally proposed in [3], in which it does not include the capacitance of the lines (a factor that can be important in LV networks). In order to define and formulate the DSSE algorithm, the following assumptions are considered:

- The network configuration is radial; it is modelled by load-buses and a slack-bus.
- Load pseudo-measures are available at every nodes of the network. Their values can be negative (injection), null (connection node), or positive (consumption). At each node, they consist of the aggregation of the end-user clients connected to that node.
- The MD, installed at a node, gives the voltage magnitude and the active/reactive power flow in the lines/transformers connected to that node.
- The lines and transformers parameters (i.e., rated values and impedances) are known.

Fig. 1 illustrates the hypotheses on the topology and the available measurements. Fig. 2 depicts the PI model of the lines that is used to formulate the Distflow equations in the presence of transversal components. The variables are computed in two steps: first, the power flow in the lines from the bottom-up is computed using the backward-Distflow equation (1); then, the square voltage magnitude of the nodes from the slack bus to the end-nodes is derived using the forward-Distflow Equation (2).

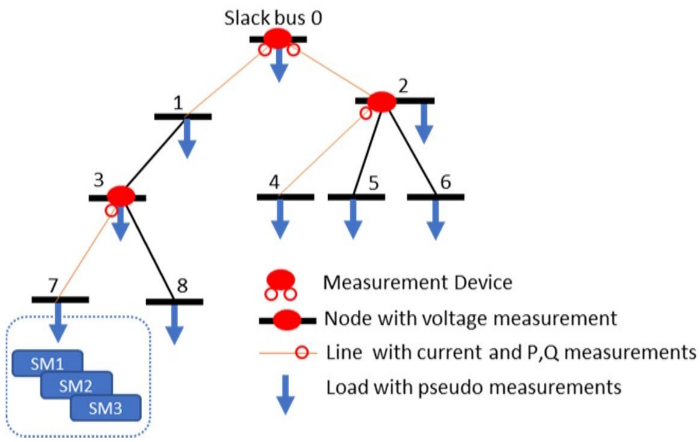

Fig. 1: Network topology example where buildings consumptions (presented by SMs) are aggregated at each node of the network.

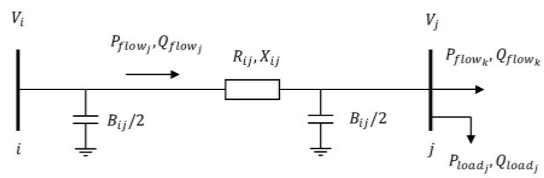

Fig. 2: PI model of a line and notations used in the Distflow.

$$\begin{cases} P_j^{flow} = P_j^{load} + \sum_k P_k^{flow} + R_j \dfrac{\left(P_j^{load} + \sum_k P_k^{flow}\right)^2 + \left(Q_j^{load} + \sum_k Q_k^{flow} + Q_{j_B}\right)^2}{V_j^2}, \\ Q_j^{flow} = Q_j^{load} + \sum_k Q_k^{flow} + Q_{j_B} + X_j \dfrac{\left(P_j^{load} + \sum_k P_k^{flow}\right)^2 + \left(Q_j^{load} + \sum_k Q_k^{flow} + Q_{j_B}\right)^2}{V_j^2}, \end{cases} \tag{1}$$

$$V_j^2 = V_i^2 - 2\left(R_j P_j^{flow} + X_j Q_j^{flow}\right) + \left(R_j^2 + X_j^2\right) \frac{P_j^{flow^2} + Q_j^{flow^2}}{V_i^2}. \tag{2}$$

As indicated in Fig. 2, P_j^{flow} and Q_j^{flow} are active/reactive power flow in line j (which is upstream to node j). P_j^{load} and Q_j^{load} are active/reactive power load at node j. V_j is the square of voltage magnitude at node j. R_j and X_j, are resistance and reactance of line j, respectively. $Q_{j_B} = -B_{\sum j} V_j^2$, where the term $B_{\sum j}$ is the sum of the shunt components (susceptances) of all the lines connected to node j.

Let us take a radial network with a slack-bus, L lines and L load-buses nodes. At each time step, the DSSE algorithm uses the weighted-least-square (WLS) estimation by iteratively solving (3).

$$\boldsymbol{z} = h(\boldsymbol{x}) + \boldsymbol{e}, \tag{3}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{z}$ is the measurement vector, $h(\boldsymbol{x})$ is the measurement function that can be derived by (1)-(2), and $\boldsymbol{e}$ is the measurement noise. Using (1) and (2), we can define the DSSE vector $\boldsymbol{x} = \left[\boldsymbol{P_{load}}\ \boldsymbol{Q_{load}}\ \boldsymbol{V_0^2}\right]^T$, where $\boldsymbol{P_{load}} = [P1 \dots P_L]^T$ and $\boldsymbol{Q_{load}} = [Q1 \dots Q_L]^T$ are the active/reactive power load of all load-buses, respectively. Finally, $\boldsymbol{V_0^2}$ is the square of voltage magnitude at slack-bus. The measurement vector $\boldsymbol{z}$ is composed of the following elements:

- From the MDs installed on the grid:
  - Voltage magnitudes at every measured node $\boldsymbol{V}^2_{\boldsymbol{n}_{mes}}$,
  - Power flow in the measured lines ($\boldsymbol{P}_{\boldsymbol{flow}_{mes}}, \boldsymbol{Q}_{\boldsymbol{flow}_{mes}}$).
- Load pseudo-measurement at every node ($\boldsymbol{P}_{\boldsymbol{load}_{mes}}, \boldsymbol{Q}_{\boldsymbol{load}_{mes}}$).

As formulated in [6], it is possible to solve the DSSE Equation (3) in an iterative manner with the help of (4).

$$\boldsymbol{x}_{(k+1)} = \boldsymbol{x}_{(k)} + (H^t W H)^{-1} H^t W \boldsymbol{r}_{(k)}, \tag{4}$$

where k is the iteration index, H is the jacobian matrix of h, W is the inverse of the variance matrix regarding all the measurements, and $\boldsymbol{r}_{(k)} = \boldsymbol{z} - h(\boldsymbol{x}_{(k)})$ is the error between the measurement and the results of DSSE at iteration k.

3 Intraday Load Forecasting Methods

As remarked in the introduction, the accuracy of DSSE highly depends on the quality of generated pseudo-measurements. Note that unlike transmission systems that enjoys high number of actual measurements [7], the distribution systems observability is undermined with a limited number of SMs and in-field MDs. In addition, the data of SMs for each day is not available in real-time and for implementing the DSSE; thus, we need to generate the pseudo-measurements based on the nodal loads profiles forecasts.

In this section, three INLF methods for generating pseudo-measurements are explained:

- SARIMA method using SMs data – Method A,
- Random forests method using SMs data – Method B,
- Random forests method using SMs and MDs data – Method C.

A: Seasonal auto-regressive integrated moving average method using SMs data

Time series models have been commonly used in a broad range of forecasting applications [8]. The SARIMA method is one of the time series models capturing serial correlation among observations both within and across the seasons.

Here, we use a separate SARIMA model for each node. To configure the SARIMA model of each node, we need the selection of the hyper-parameters for both the trend and seasonal elements of the time-series. There are the following seven elements that should be tuned: (p) trend auto-regression order, (d) trend difference order, (q) trend moving average order, (P) seasonal auto-regression order, (D) seasonal difference order, (Q) seasonal moving average order, and (m) the number of time steps for a single seasonal period. To tune these seven parameters, autocorrelation function (ACF) and partial autocorrelation function (PACF) are used/evaluated [9]. To keep the model as simple as possible and since more confident we can argue that the relationship of the recent data and the forecast is linear, we select the hyper-parameters of all SARIMA models of all nodes as $p = 1$, $q = 1$, $d = 1$, $P = 1$, $Q = 1$, $D = 1$, and $m = 24\, hours$. The performance of our selected hyper-parameters is supported by comparing the Akaike information criteria (AIC) of these models with higher order ones and testing the required time of calculating these models [9].

The final SARIMA model resulted from the above parameter selection of node n is as (4).

$$\left(1 - \phi_{n,24}.B^{24}\right).\left(1 - \phi_{n,1}.B\right).(1 - B).Y_{n,t} = \theta_{n,0} + \left(1 - \theta_{n,1}\right).\left(1 - \theta_{n,24}.B^{24}\right).\epsilon_{n,t}, \tag{4}$$

where $\phi_{n,24}$, $\phi_{n,1}$, $\theta_{n,0}$, $\theta_{n,1}$, $\theta_{n,24}$ are the parameters of the model, B is the backward shift operator, $Y_{n,t}$ is the load of node n at time t (which can be active power P_{load} or reactive power Q_{load}), and $\epsilon_{n,t}$ denotes the noise of node n load at time t. In this paper, we use the function of "SARIMAX" from the package "statsmodels" in "Python" to train/test these models.

It is worth mentioning that an SARIMA model can generally give us a good result when there are no big random changes in the time-series [10]. However, since it is not the case for the nodal loads in LV distribution networks [11], we expect that the generated pseudo-measurements by this method do not have good accuracy.

B: Random forests method using SMs data

Machine-learning methods have attracted scientific attentions for generating the pseudo-measurements that are used in DSSE [12]. One of the big challenges of using these machine-learning methods is that they need huge amounts of data and some methods are computationally expensive. In this paper, to solve this challenge, the random forests method is adopted, which has acceptable computational performance [12].

Random forests method is an ensemble learning method that can be used for both the classification and the regression purposes [6]. Random forests are the collocation of multiple decision trees known as forests. Each tree depends on an independent random sample and in the regression problem, the average of all the trees outputs is considered as the result.

In this study, we use a separate random forests model for each node n. To train each model, we need to define the input patterns $\vec{X}_n$ and output (forecast) one $\vec{Y}_n$. The forecast pattern $\vec{Y}_n$ is the vector of nodal load, i.e., $\left[Y_{n,1}\ Y_{n,2} \dots Y_{n,t}\right]^{\mathrm{T}}$. For the input pattern, we use the SMs data of the previous day; thus, $\vec{X}_n$ is

$$\begin{bmatrix} Y_{n,1-24} & \cdots & Y_{n,1-1} \\ \vdots & \ddots & \vdots \\ Y_{n,t-24} & \cdots & Y_{n,t-1} \end{bmatrix}, \tag{5}$$

where $Y_{n,t}$ is the load of node n at time t (which can be active power P_{load} or reactive power Q_{load}).

Before training the random forests model for each node, we normalize both the input and output patterns. In addition, the outliers are detected and deleted from the time-series to prevent the over-fitting. We use the function "RandomForestRegressor" in the package "sklearn" in "Python" to fit/validate these models.

C: Random forests method using SMs and MDs data

In similar to the explained method B, in method C we use random forests regression models for prediction of nodal loads and generating pseudo-measurements. The only difference is that we modify the input vector $\vec{X}_n$ by adding the data of last 24 hours of the associated MDs (in which these measurements are relevant features to the load data). Thus, the models inputs are twofold: the time-series of grid measurements and the long-term time-series of SMs measurements.

Each MD connects to a feeder including a number of nodes with SMs. For those nodes we modify $\vec{X}_n$ as

$$\begin{bmatrix} Y_{n,1-24} & \cdots & Y_{n,1-1} X_{j,1-1} & \cdots & Y_{j,1-24} \\ \vdots & \ddots & \vdots \quad \vdots & \ddots & \vdots \\ Y_{n,t-24} & \cdots & Y_{n,t-1} X_{j,t-1} & \cdots & X_{j,t-24} \end{bmatrix}, \tag{6}$$

where $X_{j,t}$ is the measurement of associated MD j at time t (which measures active power $P_{flow_{mes}}$ or reactive power $Q_{flow_{mes}}$).

It is worth mentioning that before training these models, we normalize both the input and output patterns. In addition, the outliers are detected and deleted from the time-series to prevent the over-fitting. Finally, we use the function "RandomForestRegressor" in the package "sklearn" in "Python" to fit/validate these models.

4 Test case description

In the following, a MV/LV distribution network is used to discuss the performance of different methods and highlight the characteristics of the proposed DSSE. The network and data for the case study were made available by the DSO of the city of Geneva, Switzerland. The network has 52 nodes along two feeders (Fig. 6-a). The branches connecting nodes 2 to 0 and nodes 2 to 5 are MV/LV transformer, while the others are cables. In addition, the locations of in-field MDs are depicted with red line in Fig. 6-a.

In summary, this case study network is in a residential area and has the following characteristics:

- 52 nodes, 49 lines, and 2 MV/LV transformer.
- 58 smart-meters installed (over 89 meters, so 65% of the consumer are covered):
    - Active and reactive power are available with an average over 15 minutes periods.
- 31 conventional meters, in which the annual active power data of these meters is available.
- Average monthly active power consumption of 43200 kWh.
- 2 points for MD installation:
    - 1 in the MV/LV station.
    - 1 at the start of a LV feeder.
    - Active and reactive power, three-phase voltage magnitudes, and current magnitudes are available with an average over 10 minutes periods.

In the numerical results, we compare the accuracy of generated pseudo-measurements by three explained methods A, B, and C in Section 3. In addition, for evaluating the performance of DSSE, we compare the results for the following four cases.

- Case 1: DSSE without using MDs data while the pseudo-measurements are generated by method A, i.e., SARIMA method using SMs data.
- Case 2: DSSE by using MDs data while the pseudo-measurements are generated by method A, i.e., SARIMA method using SMs data.
- Case 3: DSSE by using MDs data while the pseudo-measurements are generated by method B, i.e., random forests method using SMs data.
- Case 4: DSSE by using MDs data while the pseudo-measurements are generated by method C, i.e., random forests method using SMs and MDs data.

Here, the 15 minutes data of SMs and the 10 minutes data of MDs are acquired during 79 days from 17/01/2020 to 04/04/2020. The data of the first 71 days is used for training the models and the data of last 8 days is used for testing. In addition, the measured data of SMs and MDs is resampled into 1 hour and the numerical results of hourly forecasting and hourly DSSE are given in the following.

5 Numerical results

First, we evaluate the performance of the INLF methods for generating the pseudo-measurement as presented in section 3. It is observed that generally, Method C (i.e., random forests method using SMs data and MDs data) outperforms the other two methods. In Fig. 3, the predicted total active power and reactive power of all SMs in two mentioned feeders using method C are compared with the actual data of the last 8 days (testing data). One can observe that the sum of generated pseudo-measurements by method C is close to the actual measured data of active/reactive power.

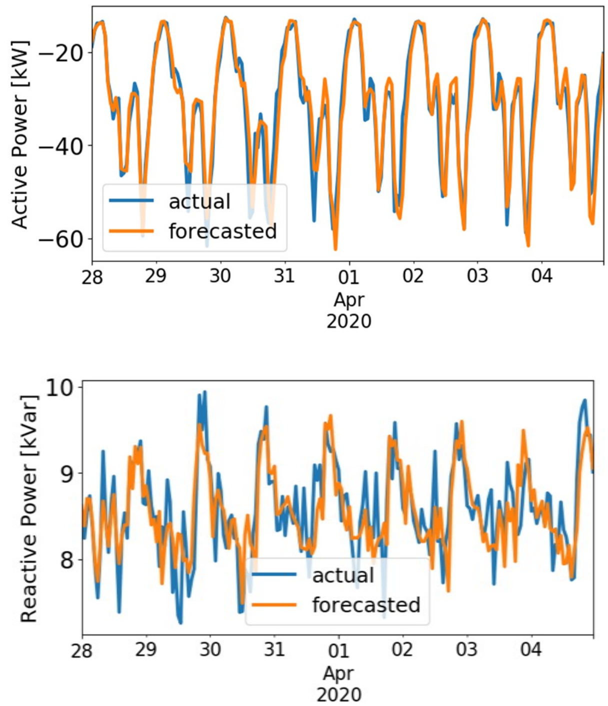

Fig. 3: Actual and forecasted total active/reactive power of all SMs using Method C.

In this study, the symmetric mean absolute percentage error (sMAPE) is used instead of mean absolute percentage error (MAPE) criteria for comparing different methods of INLF to avoid the asymmetry of the MAPE. The sMAPE is introduced by Makridakis [13] and it is more reliable than MAPE. The sMAPE for each time-series data is formulated as follows:

$$sMAPE = \frac{1}{T} \cdot \sum_{t=1}^{T} \frac{|forecasted(t) - actual\ data(t)|}{(|forecasted(t)| + |actual\ data(t)|)/2} \times 100\%. \tag{7}$$

In order to assess the performance of explained methods A, B, C in Section 3, the index of sMAPE is calculated for forecasting the active/reactive power of the nodes with SMs (16 nodes in this case study), separately. For the total active power, sMAPEs of methods A, B, and C are 14.46%, 8.8%, and 8.9%, respectively. In addition, for the reactive power, the sMAPEs of methods A, B, and C are 3.8%, 4.4%, and 3.1%, respectively. However, the sMAPEs of INLF are much higher. The box diagram of sMAPE to forecast the active/reactive power of nodal loads is depicted in Fig. 4. The averages of sMAPEs are also shown in Fig. 4. One can see that the methods B and C (random forest) outperform the method A (SARIMA) in both forecasting the active and reactive power. In addition, the worst case of sMAPEs for both active and reactive power are reduced considerably while using methods B and C instead of method A. Finally, using the MDs in forecasting (method C) has a little advantage in prediction of the nodal active power (less than 1% in average). It can be argued that each MD measures the aggregated active power of the nodes with SMs in that feeder; hence, the MDs data does not reflect the characteristics of each node profile separately. This can be improved by adding further number of MDs. In addition, for the reactive power, the measurements of MDs are even different from the aggregated reactive power of nodes with SMs since it depends on the network impedances. Thus, adding the data of MDs does not have benefit for INLF of reactive power.

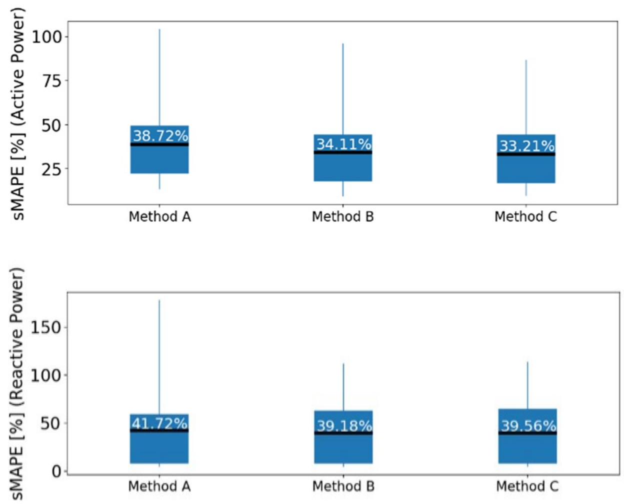

Fig. 4: Comparison of sMAPEs between different methods of A, B, and C.

The results of all four cases, explained in Section 4, are derived and compared. For comparison, the estimation error is defined as below:

$$Estimation\ Error = \frac{1}{T}.\sum_{t=1}^{T} |estimated(t) - actual\ data(t)|. \tag{8}$$

As one can see in Fig. 5, the voltage estimation error is reduced just by adding the data of MDs into the DSSE. However, for having lower current estimation error, using the random forests method is advantageous since it is conceived to have a better model of pseudo-measurements. In addition, the current estimation error of case 4 (where the data of MDs is also used in generating the pseudo-measurements) is slightly lower compared to the case 3 (where the data of SMs is just used in generating the pseudo-measurements).

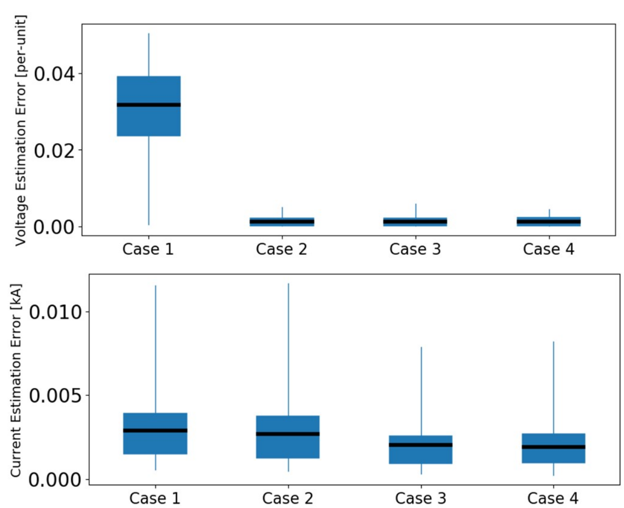

Fig. 5: Comparison of voltage and current estimation errors between different cases of 1-4.

To evaluate the geographical observability improvement of the test case study network using cases 2-4 instead of case 1, we define the improvement index for both the voltage and current estimations of each node and branch as follows:

$$\begin{aligned}&Improvement\ Index\ (Case\ j\ versus\ Case\ 1)\\&\quad = \frac{Estimation\ Error\ (Case\ j) - Estimation\ Error\ (Case\ 1)}{Estimation\ Error\ (Case\ 1)} \times 100\%.\end{aligned} \tag{9}$$

In Fig. 6-a, the locations of MDs are depicted. The improvement index of cases 2-4 versus case 1 are shown in Fig. 6-b, 6-c, and 6-d, respectively. The voltage estimations of all nodes improve more than 20% in all cases 2-4. Therefore, we conclude that the voltage estimations of all nodes can be better off just by adding MDs data into the DSSE algorithm. On the other hand, the current estimations of distant branches from the MDs locations do not improve in cases of 2-4. However, by comparing the cases 3 and 2, one can see that the current estimation of near branches improves considerably by adopting random forests method for pseudo-measurements generation. In addition, by adding the data of considered MDs in forecasting procedure and using the random forests method (case 4), the current estimations at some branches (such as 10-15, 15-19, 19-21, 21-23, 23-25, 25-29, 29-32, and 32-35) will improve. Therefore, while the advantage of case 4 compared to case 3 was not clear in Fig. 5, the performance of case 4 (using MDs data in forecasting procedure) in the DSSE current results of number of branches can be observed in Fig. 6-d.

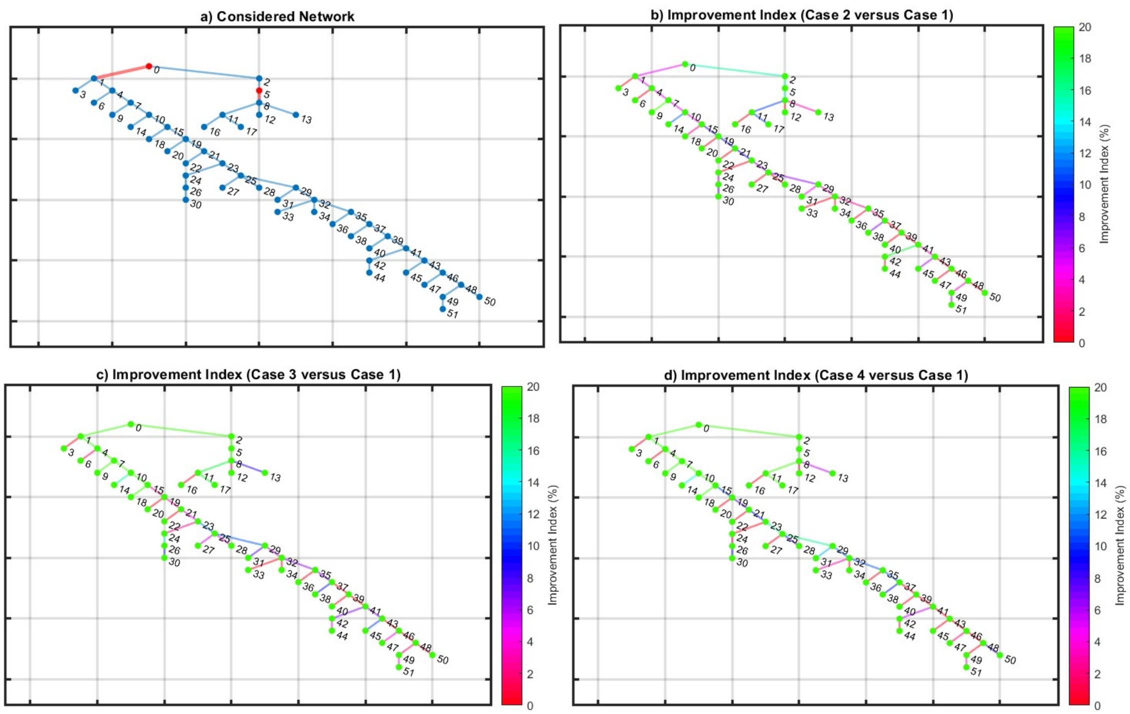

Fig. 6 a) Test case study network with MDs at red nodes/branches, b) Comparing the DSSEs of cases 2 and 1, c) Comparing the DSSEs of cases 3 and 1, and d) Comparing the DSSEs of cases 4 and 1. In figures b), c), and d), the improvement index for estimation of voltage and currents are demonstrated by color bar at each node and branch, respectively.

6 Summary and conclusions

In this paper, a distribution system state estimation (DSSE) algorithm based on Distflow model is presented. The DSSE algorithm takes the following inputs; i) near real-time data from the measurement devices (i.e., voltage magnitudes and active/reactive power flow measurements), and ii) the load pseudo-measurements obtained from an intraday nodal load forecast (INLF). The impact of forecast accuracy and available data on the DSSE results is investigated using the real data available for a distribution grid in city of Geneva, Switzerland.

It is observed that the error of INLF (30-40%) which is much higher than the error of total load forecast (4-9%). Overall, the INLF methods B and C (random forest) outperform the method A (SARIMA) in both forecasting the active and reactive power. In addition, the worst case of sMAPEs for both active and reactive power are reduced considerably while using methods B and C instead of method A. However, using MDs data in INLF (method C) only slightly improved the results (less than 1% in average), since the number of MDs are limited and the MDs data does not reflect the characteristics of each node profile separately. This can be improved by adding further number of MDs.

With reference to the numerical results, the estimation error is reduced just by adding the data of MDs into the DSSE. However, for having lower current estimation error, using the random forests method is advantageous since it is conceived to have a better model of pseudo-measurements. In addition, the current estimation error of case 4 (where the data of MDs is also used in generating the pseudo-measurements) is slightly lower compared to the case 3 (where the data of SMs is just used in generating the pseudo-measurements).

References

[1] Le Conseil fédéral Suisse., *Ordonnance sur l'approvisionnement en électricité (OApEl), section 4a Introduite par le ch. I de l'O du 1er nov. 2017, en vigueur depuis le 1er janv. 2018*. 2017.

[2] M. Carpita, A. Dassatti, M. Bozorg, J. Jaton, S. Reynaud and O. A. Mousavi, "Low Voltage Grid Monitoring and Control Enhancement: The GridEye Solution," 2019 International Conference on Clean Electrical Power (ICCEP), Otranto, Italy, 2019, pp. 94-99.

[3] A. Primadianto and C.-N. Lu, "A Review on Distribution System State Estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.

[4] H. A. R. Florez, E. M. Carreno, M. J. Rider, and J. R. S. Mantovani, "Distflow based state estimation for power distribution networks," *Energy Syst.*, Jan. 2018.

[5] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Deliv.*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.

[6] P. Paruta, T. Pidancier, M. Bozorg, and M. Carpita, "Greedy Placement of Measurement Devices on Distribution Grids based on Enhanced Distflow State Estimation", arXiv:2007.15050

[7] Abur, Ali, and Antonio Gomez Exposito. Power system state estimation: theory and implementation. CRC press, 2004.

[8] Ahn, Byung-Hoon, Hoe-Ryeon Choi, and Hong-Chul Lee. "Regional Long-term/Mid-term Load Forecasting using SARIMA in South Korea." Journal of the Korea Academia-Industrial cooperation Society 16, no. 12 (2015): 8576-8584.

[9] Contreras, Javier, Rosario Espinola, Francisco J. Nogales, and Antonio J. Conejo. "ARIMA models to predict next-day electricity prices." IEEE transactions on power systems 18, no. 3 (2003): 1014-1020.

[10] Hor, Ching-Lai, Simon J. Watson, and Shanti Majithia. "Daily load forecasting and maximum demand estimation using ARIMA and GARCH." In 2006 International Conference on Probabilistic Methods Applied to Power Systems, pp. 1-6. IEEE, 2006.

[11] Dehghanpour, Kaveh, Zhaoyu Wang, Jianhui Wang, Yuxuan Yuan, and Fankun Bu. "A survey on state estimation techniques and challenges in smart distribution systems." IEEE Transactions on Smart Grid 10, no. 2 (2018): 2312-2322.

[12] Dudek, Grzegorz. "Short-term load forecasting using random forests." In Intelligent Systems' 2014, pp. 821-828. Springer, Cham, 2015.

[13] Makridakis, Spyros. "Accuracy measures: theoretical and practical concerns." International journal of forecasting 9, no. 4 (1993): 527-529