This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

State Estimation in Power Distribution Systems Based on Ensemble Kalman Filtering

Côme Carquex, Catherine Rosenberg, *Fellow, IEEE,* and Kankar Bhattacharya, *Fellow, IEEE*

***Abstract*—In this paper, a Past-Aware State Estimation (PASE) method (for static state estimation) is proposed for power distribution systems that takes previous estimates into account to improve the accuracy of the current one, using an Ensemble Kalman Filter (EnKF). Fewer phasor measurements units (PMU) are needed to achieve the same estimation error target than snapshot-based methods. Furthermore, contrary to existing methods, the proposed approach does not embed power flow equations into the state estimator, thus making it a versatile technique. The theoretical formulation of the EnKF-based PASE presented in the paper has been validated considering a 33-bus distribution system and using power consumption traces from real households.**

***Keywords*—Distribution system, distribution system state estimation, ensemble Kalman filter, phasor measurement unit, state estimation.**

I. Introduction

STATE estimation in power distribution systems is a critical tool for ensuring a secure, reliable, and optimal performance of the system, and some utilities have already began rolling-out their implementation and use [1]. Well understood in transmission systems, static state estimation is now an area of active research in distribution networks. While several snapshot-based approaches have been used to solve this problem, only a few solutions have been proposed in a filtering-based framework; this paper focuses on static state estimation in distribution systems based on a filtering method.

The state of a power system can be completely defined from the knowledge of all bus voltage magnitudes and angles at time t [2]; typically in transmission systems, state estimation is carried out based on measurements of variables such as the voltage magnitudes and angles[1], available from Phasor Measurement Units (PMUs); power injections and power flows are commonly used as well.

At the transmission level, state estimation is traditionally carried out using a snapshot-based weighted least square (WLS) method which relies on high quality measurement data from PMUs [2]. However, transmission systems generally have a limited number of buses and are equipped with many measurement devices since it is important to precisely monitor and control the system at all times. On the other hand, distribution systems comprise a large number of buses with little measurements available. While PMUs are not yet widely available at the distribution level, it is expected that they will become more prevalent in the future. Indeed several recent studies have focused on developing low-cost, easy to deploy PMUs [3], [4]. It is nonetheless not practical to install PMUs at every distribution bus. If PMUs were to be placed at selected buses only, there would be infinitely many solutions to the DSSE problem. In order to reduce the number of possible solutions, pseudo-measurements can be used [5], which are load forecasts computed ahead of time to aid DSSE in finding a "good" solution. Typically, a pseudo-measurement at a given load bus comprises an estimate of the expected active and reactive power consumptions at the bus. Load forecasting at the distribution level is difficult, hence pseudo-measurements are usually of poor quality. These fundamental differences between transmission system state estimation and DSSE, and the need for affordable solutions, mean that new state estimation approaches are needed for distribution systems.

Many studies have extended the WLS approach from transmission to distribution systems. A review of literature on the different state estimation techniques and their application to DSSE problems is presented in [6]. One of the first applications of the snapshot approach to the DSSE problem was reported in [7], where a probabilistic formulation based on pseudo-measurements was used. In [8], the power-flow equations were linearized and a computationally friendly solution method was proposed. The authors also showed that PMUs are needed for accurate state estimation. Compressed sensing theory was used for state estimation with sparse measurements in [9], while [10] used line-current magnitudes and angles. Finally a semi-definite programming approach was used to solve the DSSE problem in [11].

Several researchers have used Kalman filters in state estimation problems for transmission systems [12]. However, in distribution systems, the poor quality of the pseudo-measurements renders such methods ineffective. Therefore, very few Kalman filtering based methods have been developed for DSSE and none improve over the WLS. Huang *et al*. compared the extended Kalman filter to the unscented Kalman filter in [13]. From the reported results it was noted that there was no visible improvement in performance of the Kalman filter based methods over WLS. In [14] the impact of choice of the model and measurement covariance matrix on the performance of the extended Kalman filter was examined. It was noted from the results that the proposed filtering approach did not result in any performance improvement. The above discussed Kalman filter based approaches apply the methods directly from the transmission to distribution systems. The problem of poor quality of pseudo-measurements is alleviated by assuming that measurements are available at every bus in real-time or quasi-real-time, usually from synchronized smart-meters, which is

C. Carquex, C. Rosenberg and K. Bhattacharya are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada, e-mail: {cacarque, cath, kankar}@uwaterloo.ca

[1]While phase angles are generally available, not all utilities have updated their state estimation software to use them.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

not realistic[2].

In a snapshot-based context where the state at time t is computed independently of the estimates at times anterior to t and where the measurement errors are independent, identically distributed and follow a Gaussian distribution, the WLS objective function provides the best performance possible (excluding ill-conditioned cases) [15]. Such an estimator is referred to as the State of the Art (SoA) in this paper, for the purpose of comparison.

In this paper, a past-aware method for DSSE, named PASE (Past-Aware State Estimation), where the estimate at time t depends on anterior estimates and based on the Ensemble Kalman Filter (EnKF) [16] is presented. Applying the EnKF to this problem is non-trivial, since measurements from sources with different time-scales must be merged. Contrary to WLS and other approaches using different variations of the Kalman filter, the proposed PASE approach does not embed the power flow equations into the estimator, making it a versatile technique. Instead it relies on an external power-flow solver, which is left to the choice of the operator. This paper focuses on overcoming the challenges related to filtered state estimation and analyzing when using a filtered approach is beneficial. An analytical method for estimating the performance gain brought about by PASE over the SoA in a time-efficient way is introduced as well, which reduces the need for expensive Monte-Carlo simulations.

In view of the above discussions, the main contributions of this work are:

- A maiden attempt is made to apply EnKF to a power distribution system sparsely monitored by PMUs for state estimation.
- An analytical framework is developed to evaluate the performance of PASE.
- The theoretical results are validated via extensive simulations on a 33-bus distribution system [17] using power consumption traces from real households.
- The performances of the proposed PASE approach and WLS are compared and engineering insights are presented to understand the impact of each decision variable on the performance of PASE, as well as the trade-offs to make.

Based on the above discussions, the main message of this work is that PASE is the first technique to improve upon the SoA. It does so significantly when the elapsed time between two consecutive state computations is small (less than 15 minutes), i.e., less PMUs are needed to achieve the same estimation error.

The rest of the paper is organized as follows. The background and assumptions are presented in Section II. The SoA method is presented in Section III and the proposed PASE solution in Section IV. The validation results are reported in Section V. Finally, the conclusions are drawn in Section VI.

[2]Indeed currently available smart-meters typically report energy consumptions once a day, at the end of the day, with a typical granularity of 15 minutes (other granularities such as 5, 30 or 60 minutes for example can also be found). Making the smart meters report their measurements more often (e.g., every 15 minutes) will impact their cost.

II. System and Assumptions

In order to demonstrate the performance of the proposed DSSE method, the distribution system is assumed to be three-phase balanced, and operating under normal conditions. Also that, the DSSE problem is being solved by the local distribution company (LDC) using an appropriate computational platform. The following information are necessary in order to implement the DSSE, both with the SoA method and the proposed PASE method.

Computational timescale: A new state estimate is computed every ΔT. Typically in transmission systems, time-step of 1 min or less are considered [18], [19]. However, in distribution systems smaller time-steps are needed because of higher load volatility, which can arise, for example, with high penetration of renewables. The value of ΔT has an impact on the computational burden. In this work, time-steps from 6 seconds to 15 minutes are considered. Nevertheless, the choice of an appropriate timescale for DSSE problems is still an open question.

Topology: The distribution system has a radial topology and is defined by a set of buses I of cardinality $|I|$ as well as a set of branches B of constant and known impedances, connecting the buses. While in this work the network model is assumed to be perfectly known, in practice, the precision of the model would impact the accuracy of the estimated state, irrespective of whether PASE or any other method is used. This assumption is however commonly made in almost every state estimation work. The substation transformer is modeled as a reference voltage source of magnitude V_0.

Measurements: The subset $S \subseteq I$ of buses are equipped with PMUs that monitor at every ΔT both, bus voltage magnitudes (V_s) and bus angles (δ_s). In this paper, the PMUs are assumed to provide only the voltage phasors and not the current phasors of the branches since the focus of this work is on low-cost PMUs [3], [4], as mentioned in the introduction (clearly PASE and the SoA can accommodate other types of measurements (such as branch flows for example)). A broadband communication infrastructure is available to transmit the measurements with low latency and high reliability. The PMUs are placed in the distribution system according to a given mapping $\mathscr{S}$.

Pseudo-measurements: these are forecasts that "measure" both active and reactive powers. They are available for each bus i in I. Forecasts are made at periodic intervals $\Delta T'$, typically once a day for the next day (day-ahead forecast). At the time of computation, the most recent forecast is used. Clearly, forecasts and PMU measurements are on completely different time-scales ($\Delta T' \gg \Delta T$), hence the non-triviality of the EnKF. Forecasts are made based on historical data. Previous estimation work based on Kalman filters assumed real-time consumption data. This strong requirement is relaxed with forecasts. This time horizon is well suited for DSSE since day-head forecasts are typically computed by utilities every day. Hence they are readily available and do not introduce any extra computational burden [20].

Data requirements: both the SoA and PASE approaches require a forecasting method as well as sample power con-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

sumption traces (active and reactive) from the system at the level of each distribution transformer, from which the forecasting method can be calibrated. Using the data, error parameters can be obtained offline. Let $e_i(t)$ be the forecast error at bus i and time t (for active power, for example); $e_i(t)$ is assumed to be a stationary random process. The variance of the forecast error ($\mathrm{E}[e_i(t)^2]$) is supposed to be known. These two hypotheses are almost always used by researchers [8]. The estimation of the variance of the forecast errors comes from the acquired data.

The proposed PASE method needs two additional information that can be derived from the same sample data: a load evolution model (which will be discussed in Section IV-A) and the forecast error correlation coefficient, evaluated between two (computation) time-steps at a given bus (i.e., $\mathrm{E}[e_i(t)e_i(t-\Delta T)]$).

Finally, the load forecast errors are assumed to be uncorrelated between buses, an assumption often made in the literature [8][3].

System state: it is represented by state vectors; different (equivalent) state representations are used depending on their ease of use in the problem formulation. For example,

$$\mathbf{y}[t] = [\mathbf{V}[t]^T, \boldsymbol{\delta}[t]^T]^T$$

is a possible state vector representation, where $\mathbf{V}[t]$ is the vector of voltage magnitudes at each bus, and $\boldsymbol{\delta}[t]$ the vector of voltage angles. Another way is to define $\mathbf{x}[t] = [\mathbf{P}[t]^T, \mathbf{Q}[t]^T]^T$ where $\mathbf{P}[t]$ and $\mathbf{Q}[t]$ denotes the vectors of active and reactive power injections at each bus, respectively [21]. Note that given that the substation transformer is modeled as a constant voltage source, the source bus voltage is not included in the state vector. Also note that the power-flow equations link the state-vectors $\mathbf{x}$ and $\mathbf{y}$. A third way, used in theoretical formulations, is $\mathbf{w}[t] = [\underline{v}_1[t], \ldots, \underline{v}_{|I|}[t]]^T$ where $\underline{v}_i[t]$ is the voltage phasor at bus i, time t; this can also similarly be related to other representations.

Limitations: In this work, unbalanced system, distributed generation and biased measurements are not considered and are left for future studies.

III. State-of-the-Art DSSE Method

The SoA method [2], [22], [23] used to solve the DSSE problem, is a snapshot approach and uses a nonlinear WLS objective function. The inputs and outputs of the SoA are summarized in Fig. 1. Given the system characterized by the sets I, B, S and the mapping $\mathscr{S}$, the system state, at a given time, is estimated using an overdetermined set of equations. In the following, the time dependency of the variables is dropped for better readability. The variables to be determined are the $2|I|$ state variables. Each measurement adds one constraint. There are either 2 or 4 measurements per bus (active/reactive power forecast, voltage magnitude, and angle), depending on whether there is a PMU at the bus. The number of constraints is given by $M = 2|I| + 2|S|$.

[3]Note that even if this assumption is not made, the SoA and PASE formulations are still valid. In such a case, the correlations need to be taken into account in the computations.

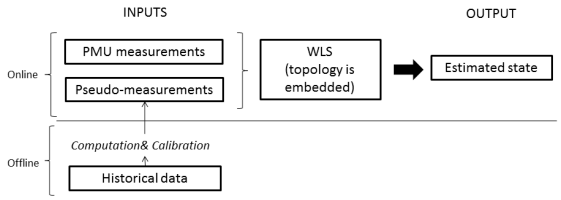


Fig. 1. Flowchart detailing the inputs and output of the SoA

The PMU measurements and the forecasts are stored in a vector $\mathbf{z}$ of length M, and are related to the system state as per the following model: $\mathbf{z} = f(\mathbf{y}) + \boldsymbol{\eta}$ where f is the function that maps the state vector to the measurement vector, and η is the vector containing the measurement noise and model uncertainties. For example, $f(y) = [\mathbf{V}(y)^T, \boldsymbol{\delta}(y)^T, \mathbf{P}(y)^T, \mathbf{Q}(y)^T]^T$ where $\mathbf{V}(y)$ and $\boldsymbol{\delta}(y)$ are the vectors, respectively, containing the voltage magnitude and angle measurements at the buses with PMUs and $\mathbf{P}(y)$ and $\mathbf{Q}(y)$ are vectors of active and reactive power forecasts of size $|I|$, respectively. Assuming that the measurement errors are uncorrelated and have zero mean, the covariance matrix Σ of the error vector $\boldsymbol{\eta}$ is written as, $\Sigma = \mathrm{diag}(\sigma_1^2, ..., \sigma_M^2)$, where σ_m^2 is the variance of the m^{th} measurement.

The objective function to be minimized at each time-step is given below:

$$J(\mathbf{y}) = (\mathbf{z} - \boldsymbol{f}(\mathbf{y}))^T \Sigma^{-1} (\mathbf{z} - \boldsymbol{f}(\mathbf{y})) \tag{1}$$

Several methods exist to minimize the objective function, the simplest being to iteratively linearize f and solve the resulting objective using the normal equations.

IV. Proposed Method: PASE

To solve the DSSE problem, PASE, an EnKF-based method, is proposed. Kalman filters are sequential filtering methods. Each iteration is a two steps process: 1) the system state is integrated in time using an evolution model, defining a (*a priori*) state estimate. 2) Available measurements (including pseudo-measurements) are used to correct the estimate and define the updated state. The second step is referred to as the *update-step* during which data *assimilation* occurs. The load evolution model used in this approach is presented in Section IV-A. The idea behind the proposed approach is simple: the additional information provided by the load evolution model and the previously estimated states are used to alleviate the poor quality of pseudo-measurements. The inputs and outputs of PASE are summarized in Fig. 2. Notice how the flowchart for PASE is very similar to the one for the SoA.

A. Load Evolution Model

For each distribution transformer bus, an evolution model for the aggregate load is needed, both for the active and reactive power consumptions. An autoregressive model of order one AR(1) is used, with coefficient equal to one. Such a model is chosen for the following main reasons: it is simple, fits within the Kalman filter framework and is intuitively reasonable, for the time horizon considered in this work. In addition, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

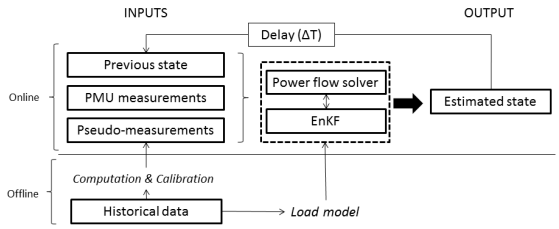


Fig. 2. Flowchart detailing the inputs and output of PASE

Algorithm 1 Estimation of the state at time-step k

Input: X^{k-1}, measurements and pseudo-measurements at time-step k.

1: Compute X_p^k: integrate the ensemble in time (Eq. 3)
2: Compute X_u^k: assimilate pseudo-measurements (Eq. 11)
3: Compute X_a^k: assimilate PMU measurements (Eq. 13)
4: $X^k \leftarrow X_u^k$

Output: Estimated state $\tilde{\mathbf{x}}^{\mathbf{k}} = E[X^k]$ for time-step k.

EnKF requires the characterization of the prediction error which is easily done with such a model. IT is discussed next. Let $L_i^p(t)$ and $L_i^q(t)$ denote the instantaneous active and reactive aggregated power respectively, at bus i and time t. It is assumed that $L_i^p(t)$ and $L_i^q(t)$ are stationary random processes [24]–[26]. The representations of L_i^p and L_i^q are such that: $L_i^p(t+\Delta T) = L_i^p(t) + \zeta_i^p(t)$, where $\zeta_i^p(t)$ is white noise (and similarly $\zeta_i^q(t)$ for L_i^q). The load evolution model will characterize the probability density functions (pdf) of $\zeta_i^p(t)$ and $\zeta_i^q(t)$. Specifically, the load variation between two (computation) time-steps is considered. The load variation (aka load evolution model) for active and reactive powers are defined as the stationary random processes $L_i^p(t) - L_i^p(t-\Delta T)$ and $L_i^q(t) - L_i^q(t-\Delta T)$ respectively, characterized by their probability density functions (pdf). The mean of the processes is zero and the variance of the processes can be computed from the pdf both for active and reactive powers at bus i, denoted by $(\sigma_i^p)^2$ and $(\sigma_i^q)^2$, respectively. Such an evolution model is simple and fits within the EnKF framework. The pdf can be derived empirically, for example, from the existing required sample traces, discussed in Section II as will be explained in Section V-A. The empirical pdfs are computed once, offline, for each bus. Clearly a given load evolution model is valid only for systems with similar load compositions, and will vary for different geographical areas.

B. Ensemble Kalman Filter

The traditional Kalman filter maintains a covariance matrix associated with the state estimate. The EnKF does not use such a matrix and represents the system state pdf using a set of state vectors called ensemble. Such ensemble at time-step k (i.e., time $k\Delta T$) is named X^k. The covariance matrix is replaced by the empirical covariance computed from the ensemble. The estimated system state is simply the mean of the ensemble columns. The size of the ensemble, L, will impact performance. A small ensemble size will yield faster computations. However the covariance estimate from the ensemble will be less accurate. Therefore there is a trade-off between computational speed and accuracy and a typical choice is L between 500 and 1000 [16]. The ensemble size L is independent of the state vector size. The covariance estimator $\mathrm{cov}(A,B)$ of two ensembles A, B is defined as [16]:

$$\mathrm{cov}(A,B) = \frac{1}{L-1}(A - \mathrm{E}[A])(B - \mathrm{E}[B])^T \qquad (2)$$

where $\mathrm{E}[A]$ is the estimator of the mean of the column vectors contained in ensemble A. For $\mathrm{cov}(A,A)$ the shorter syntax $\mathrm{cov}(A)$ is used. Each iteration of the EnKF (corresponding to a computation of the state vector at time-step k) follows the procedure detailed in Algorithm 1, each step of the algorithm is discussed next.

C. Initial Ensemble

The state vector $\mathbf{x} = [\mathbf{P}^T, \mathbf{Q}^T]^T$ (of size $2|I|$) is used. It is chosen given that the load evolution model described in Section IV-A is defined in terms of injected power. The pdf of the state vector $\mathbf{x}$ is represented by an ensemble of size L: $X^0 = [\mathbf{x}_1^0, \ldots, \mathbf{x}_L^0]$, X^0 is a $2|I| \times L$ matrix containing the ensemble members. The initial ensemble is built by choosing a "best-guess" estimate $\mathbf{x}^0$ of the state vector, to which perturbations are added to represent the error statistics of the initial guess. The choice of the initial ensemble is discussed in Section V.

D. Ensemble Integration

The EnKF is considered at time-step k. The prior ensemble X_p^k is obtained by individually integrating forward in time each vector of the ensemble X^{k-1}, which was computed at the previous time-step. Given the AR(1) process considered, the integration is such that:

$$X_p^k = X^{k-1} + [\boldsymbol{n}_1, \ldots, \boldsymbol{n}_L] \qquad (3)$$

where $\boldsymbol{n}_l$ $(l = 1, \ldots, L)$ are column vectors of size $2|I|$ containing the stochastic noise which accounts for the uncertainties of the load evolution model. Based on the load evolution model defined in Section IV-A, two variance values $(\sigma_i^p)^2$ and $(\sigma_i^q)^2$ are associated to each bus i $(i = 1, \ldots, |I|)$, respectively for the active and reactive powers. Their values depend on the empirical pdf derived.

Each $n_{i,l}$ and $n_{|I|+i,l}$ $(i = 1, \ldots, |I|)$ is respectively drawn from a distribution which represents the empirical pdf of the load evolution model. Note that the EnKF can accept any load evolution model.

E. Assimilation of Pseudo-Measurements

The assimilation of measurements and pseudo-measurements correspond to the update step of the Kalman filter, described at the beginning of Section IV.

An assumption in Kalman filtering is that the measurement error is white Gaussian noise. Since pseudo-measurements are forecasts and do not depend on the state of the system, they do not satisfy this requirement; instead the forecast error is correlated in time. This problem, which is recurrent in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

Kalman-based kinematic GPS applications has been solved previously, and a summary of the different existing techniques can be found in [27]. The solution chosen in this paper is the time-differencing approach described in [28] to remove time-correlated error in the pseudo-measurements. This method was selected for two reasons: 1) it does not require any reinterpretation of the Kalman equations and 2) it does not introduce any latency.

To remove the correlations, the following process is used. Let the transition matrix Ψ of the time-correlated error be defined as:

$$\Psi = \text{diag}(\psi_1^p, \ldots, \psi_{|I|}^p, \psi_1^q, \ldots, \psi_{|I|}^q) \tag{4}$$

where ψ_i^p and ψ_i^q $(i = 1, \ldots, |I|)$ are the forecast error correlation coefficients at bus i, respectively for active and reactive powers, introduced in Section II;Ψ is diagonal since the forecast errors between buses are assumed to be uncorrelated. Q is defined as the model noise covariance matrix, and is given as:

$$Q = \text{diag}((\sigma_1^p)^2, \ldots, (\sigma_{|I|}^p)^2, (\sigma_{|I|+1}^q)^2, \ldots, (\sigma_{2|I|}^q)^2) \tag{5}$$

R is the covariance matrix of the forecast error, of size $2|I| \times 2|I|$. R is diagonal since the forecast errors are assumed not correlated across buses, and is given as:

$$R = \text{diag}((\sigma_1^{fp})^2, \ldots, (\sigma_{|I|}^{fp})^2, (\sigma_{|I|+1}^{fq})^2, \ldots, (\sigma_{2|I|}^{fq})^2) \tag{6}$$

where σ_i^{fp} and σ_i^{fq} are the standard deviations of the forecast error at bus i, respectively for the active and reactive powers. The pseudo measurements are contained in a vector $\mathbf{d}$ of size $2|I|$. An ensemble D of L perturbed observations is defined such that $D = [\mathbf{d}_1, \ldots, \mathbf{d}_L]$ with each $\mathbf{d}_l = \mathbf{d} + \boldsymbol{\epsilon}_l$ $(l = 1, \ldots, L)$, where $\boldsymbol{\epsilon}_l$ is a vector drawn from a distribution which models the pseudo-measurement noise. Before establishing the update step, intermediary matrices are defined next, which will be reused for the theoretical derivations.

$$H^{\triangleright} = H - \Psi H, \quad C = QH^T\Psi^T, \quad D^{\triangleright} = D - \Psi D \tag{7}$$

$$R^{\triangleright} = (R - \Psi R\Psi^T) + \Psi HQH^T\Psi^T \tag{8}$$

The updated observation and measurement matrices, $H^{\triangleright}$ and $D^{\triangleright}$, respectively, are computed in (7). The updated measurement error matrix $R^{\triangleright}$ is computed in (8); Ψ is used to remove the time correlation of the forecast error between two time-steps; C is used as intermediary to shorten (9)-(10). The model noise matrix Q is needed to ensure that the noise introduced by the evolution step is retained. Indeed such noise does not have any time correlation component. In this context, the observation matrix H is the identity matrix (in Section IV-G the observation matrix will not be the same). The update equations for the assimilation of pseudo-measurements are given as (E is an intermediary matrix, K represents the Kalman gain matrix):

$$E = H^{\triangleright}\text{cov}(X_p^k)H^{\triangleright T} + R^{\triangleright} + H^{\triangleright}C + C^T H^{\triangleright T} \tag{9}$$

$$K = (\text{cov}(X_p^k)H^{\triangleright T} + C)E^{-1} \tag{10}$$

$$X_u^k = X_p^k + K(D^{\triangleright} - H^{\triangleright}X_p^k) \tag{11}$$

F. Assimilation of PMU Measurements

Similar to the pseudo-measurements, the measurements coming from the PMUs are contained in a vector $\mathbf{z}$ of size $2|S|$. An ensemble Z of L perturbed observation vectors is computed such that $Z = [\mathbf{z}_1, \ldots, \mathbf{z}_L]$, with each $\mathbf{z}_l = \mathbf{z} + \boldsymbol{\xi}_l$ $(l = 1, \ldots, L)$, where $\boldsymbol{\xi}_k$ is a vector drawn from a distribution which models the measurement noise.

The measurements from the PMUs can be related to the state vector using a function h, such that $\mathbf{z}_l = h(\mathbf{x}_l) + \boldsymbol{\gamma}_k$, where $\boldsymbol{\gamma}_k$ is an error vector. The function $h(\cdot)$ takes as input the system state and returns a vector containing the measurements that would have been observed considering that particular system state. Given that $\mathbf{x}$ contains the active and reactive powers injected at each bus, $\boldsymbol{h}(\cdot)$ is the power-flow solution; the EnKF does not need to know the analytical expression of $\boldsymbol{h}(\cdot)$. It is the solution given by the LDC's power-flow solver, for example. This makes the EnKF independent of the way power-flows are computed. The cost of such independence is computational: one need to compute L power-flows at each time-step. Since $\boldsymbol{h}(\cdot)$ is non-linear, the measurements cannot be obtained directly from the state using a simple multiplication by an observation matrix. Instead, $\boldsymbol{h}(\mathbf{x})$ needs to be computed explicitly. A temporary augmented state $\widehat{\mathbf{x}}$ and augmented ensemble $\widehat{X}_u^k$ are used to perform the assimilation, where:

$$\widehat{\mathbf{x}}_l = [\mathbf{x}_l^T, \boldsymbol{h}^T(\mathbf{x}_l)]^T, \quad \widehat{X}_u^k = [\widehat{\mathbf{x}}_1, \ldots, \widehat{\mathbf{x}}_L] \tag{12}$$

The updated ensemble X_a^k is then computed:

$$X_a^k = X_u^k + K(Z - \widehat{H}\widehat{X}_u^k) \tag{13}$$

$$K = \text{cov}(X_u^k, \widehat{H}\widehat{X}_u^k)[\text{cov}(\widehat{H}\widehat{X}_u^k) + \text{cov}(Z)]^{-1} \tag{14}$$

where $\widehat{H}$ is a selection matrix used to select the rows of the state vector corresponding to the desired measurements. While only PMU measurements are considered in this paper, other types can be assimilated using the same technique.

G. Theoretical Estimate of Performance

In this section, a method to compute a theoretical estimate of the performance and the improvement achieved by the proposed PASE method is developed. The goal of this theoretical modeling is to be able to check in a time-efficient manner if the performance gain brought about by PASE is worth its higher complexity, on the system considered. Indeed, computing the performance gain by running Monte-Carlo simulations is costly in time and computations. Therefore, being able to have a rough estimate of the performance gain is extremely useful. The method is based on [8], where the authors proposed a technique for estimating a priori the performances of the WLS estimator. Their work is extended in this paper to fit the EnKF and compute the relative gain between the two. The derivation is performed under the following assumptions, also made in [8]. The state vector is represented by $\mathbf{w} = [\underline{v}_1, \ldots, \underline{v}_{|I|}]^T$. The forecast variance, the forecast error time-correlation and load evolution model variance are assumed to be constant and identical for active and reactive powers. They are denoted respectively $(\sigma_i^f)^2$, ψ_i^f and $(\sigma_i^d)^2$. At each bus i, the apparent power magnitude $|S_i^f|$ is used to represent the load forecast. In

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

the analysis framework, the shape of the load evolution model is not need to be known, the value of the variance is sufficient.

The theoretical computations are performed by using a linear Kalman filter. The covariance matrices are made time-invariant in order to obtain a steady-state formulation of the filter [29]. From this formulation, the covariance matrix of the system state can be computed and used to approximate the performance of the non-linear EnKF. The performance of WLS can also be computed since it can be seen as a Kalman filter that is reset for each new estimation. To evaluate the performance of the two state estimators over a period of time T, the average root mean square error of the voltage estimate (ARMSEV) is used as metric:

$$\text{ARMSEV} = \sqrt{\frac{1}{T|I|}\sum_{t=0}^{T}\sum_{i=1}^{|I|}\mathbf{E}[|\hat{\underline{v}}_i[t] - \underline{v}_i[t]|^2]} \tag{15}$$

where $\underline{v}_i$ is the true voltage at bus i and $\hat{\underline{v}}_i$ the estimated one. A linear version of the power-flow equations is used; it is the first iteration of backward-forward sweep. A vectorized formulation is obtained by using a distribution load flow (DLF) matrix, denoted by M, as described in [30]. The relationship between the injected power at each bus (represented by the vector $\mathbf{s} = [\underline{s}_1, \ldots, \underline{s}_{|I|}]^T$, with $\underline{s}_i$ the injected power at bus i) and the state vector is given as:

$$\mathbf{w} = [V_0, \ldots, V_0] + \frac{1}{V_0} M \cdot \bar{\mathbf{s}} \tag{16}$$

where $\bar{\mathbf{s}}$ is the conjugate of $\mathbf{s}$. Several matrices used by the Kalman equations are defined. The load evolution noise covariance matrix Q expressed in terms of the apparent power, and the forecast error covariance matrix R_S are computed as follows:

$$Q = \text{diag}((\sigma_1^d)^2, \ldots, (\sigma_{|I|}^d)^2) \tag{17}$$

$$R_S = \text{diag}((\sigma_1^f)^2, \ldots, (\sigma_{|I|}^f)^2) \tag{18}$$

The PMU measurement error covariance matrix is approximated by assuming that the variance of the voltage error when projected onto the real and imaginary axes is the same and equal to $\sigma_{PMU}^2 V_0^2$, where σ_{PMU}^2 is the relative variance of the PMU measurements such that $R_{PMU} = 2\sigma_{PMU}^2 V_0^2 \times I_{|S|}$, where $I_{|S|}$ is the $|S| \times |S|$ identity matrix.

The steady state covariance matrix of the state vector is computed by iterating the Kalman equations. The covariance matrix is denoted by $\Sigma_a^{(\cdot)}$. The iteration number is indicated in the parenthesis $(\cdot)$. Such matrix will converge to a steady state covariance matrix $\Sigma_a^{(ss)}$. For each iteration, two other matrices are used to track the covariance matrix during intermediary steps: $\Sigma_p^{(\cdot)}$ and $\Sigma_u^{(\cdot)}$. They represent respectively the covariance matrix of the prior state and the state after assimilation of PMU measurements. At iteration 0, the prior covariance matrix of the state is computed such that:

$$\Sigma_p^{(0)} = M \cdot R_S \cdot M^H \tag{19}$$

where $(\cdot)^H$ indicates the Hermitian transpose (transpose conjugate operator). The updated covariance matrix obtained after the assimilation of the PMU measurements is then computed:

$$\Sigma_a^{(0)} = \Sigma_p^{(0)} - KH\Sigma_p^{(0)} \tag{20}$$

$$K = \Sigma_p^{(0)} H^T (H\Sigma_p^{(0)} H^T + R)^{-1} \tag{21}$$

where H is the observation matrix for PMU measurements. It is a selection matrix that relates state variables to the measurement vector. One can estimate the ARMSEV performance of WLS based on $\Sigma_a^{(0)}$: $\text{ARMSEV}_{\text{WLS}} = \sqrt{\frac{1}{|I|}\text{trace}(\Sigma_a^{(0)})}$.

Any iteration it ($it \neq 0$) is performed in 3 steps: first the prior covariance matrix $\Sigma_p^{(it)}$ is computed based on the previous iteration, then the covariance matrix is updated using the PMU measurement covariance matrix, and finally the pseudo-measurements are assimilated. The first two steps are such that (where H is the same matrix as in (20)):

$$\Sigma_p^{(it)} = \Sigma_a^{(it-1)} + M \cdot Q \cdot M^H \tag{22}$$

$$\Sigma_a^{(it)} = \Sigma_p^{(it)} - KH\Sigma_p^{(it)} \tag{23}$$

$$K = \Sigma_p^{(it)} H^T (H\Sigma_p^{(it)} H^T + R_{PMU})^{-1} \tag{24}$$

The third step differs because of the pseudo-measurement error time correlation (see Section IV-E). The same time-differentiation method is used. The same updated matrices are computed according to (7)-(8) with only a few differences. Now H is the inverse DLF matrix, mapping the state vector to the injected power ($H = M^{-1}$). The forecast error time correlation matrix is such that $\Psi = \text{diag}(\psi_1^f, \ldots, \psi_{|I|}^f)$. Finally, the matrix R used in (8) is such that $R = R_S$. The update equations thus become:

$$\Sigma_a^{(it)} = \Sigma_u^{(it)} - (\Sigma_u^{(it)}(H^\triangleright)^T + C) \cdot K^T \tag{25}$$

$$K = \begin{array}{l}[\Sigma_u^{(it)} \cdot (H^\triangleright)^T + C] \cdot [H^\triangleright \Sigma_u^{(it)} (H^\triangleright)^T \\ \quad + R^\triangleright + H^\triangleright C + C^T (H^\triangleright)^T]^{-1}\end{array} \tag{26}$$

Once the steady state is reached after a few iterations, the theoretical performance of the EnKF can be computed. The ARMSEV error is such that: $\text{ARMSEV}_{\text{EnKF}} = \sqrt{\frac{1}{|I|}\text{trace}(\Sigma_a^{(ss)})}$. The relative gain is expressed as: $\text{Gain} = \frac{\text{ARMSEV}_{\text{WLS}} - \text{ARMSEV}_{\text{EnKF}}}{\text{ARMSEV}_{\text{WLS}}}$.

V. Validation and Results

The improvement in performance achieved by the proposed PASE method over WLS is evaluated by considering a 33-bus test distribution feeder [17] under normal operations. Its one-line diagram is given in Fig. 3. The WLS estimation problem is modeled in GAMS environment and solved using the MINOS solver. Attention has been paid to avoid potential numerical issues. The ensemble size is set to $L = 500$ and the power flow solutions obtained from $h(\cdot)$ using the backward/forward sweep method [30]. The system is simulated over a period of 24 hours. For the theoretical estimation, 50 iterations ($(ss) = 50$) are enough to compute the steady-state of the state covariance matrix.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

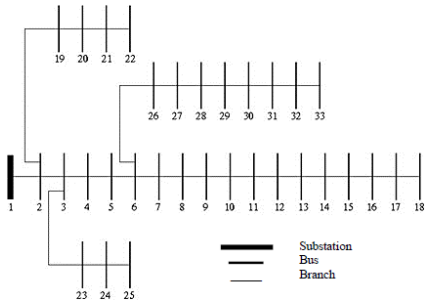


Fig. 3. One-line diagram of the 33-bus system.

A. Load Evolution Model

The (bus) load evolution model was developed using a fine-grained energy consumption dataset from Ontario, Canada. The dataset used to build the model is described in [31] and comprises instantaneous active power consumption data from 20 homes, collected over eight months, with a resolution of 6 seconds. The dataset is split randomly into two subsets, one for deriving the characterization (training set), and one for the validation process (testing set). No distinction is made between the size of the houses nor for special days. The resulting dataset is a collection of a few thousands of traces. Although 20 homes may seem to be a limited sample size, considering the daily power traces independently allows to have a large number of unique profiles. Moreover the 20 households cover a wide range of living area sizes and energy consumption patterns which increases the trace diversity.

Let n be the number of households connected to a bus. Using the training set, empirical distributions for load changes were constructed for different values of time-steps ΔT and aggregation levels n. A Laplace distribution described by a scale parameter b (and variance $\sigma^2 = 2b^2$) was found to be a good fit. The mean value is set to zero since as many positive and negative load changes are expected. This implies that the transition model is the identity, while its uncertainty is characterized by the Laplace distribution. Figures 5 and 6 are Q-Q plots that illustrate respectively, a good fit and the worst fit of all the aggregation levels and time-steps, considered in Fig. 4. In Fig. 5, the Q-Q plot approximately lies on a straight line. In Fig. 6 (the worst fit), the S shape indicates that the empirical distribution has a lighter tail than the Laplace fit. The light tail property is desirable as it means that the load evolution model will be conservative in its estimate of the uncertainty.

The influence of ΔT and n on the distribution variance is illustrated in Fig. 4. The variance essentially describes the load variation over time, a small value implying little variations. It is noted that as n decreases and ΔT shrinks, the value of σ^2 diminishes.

It was assumed that load changes are uncorrelated between buses; which can be verified to hold true from the dataset, for any value of n and ΔT up to 30 minutes.

The values of σ^2 are derived empirically as a function of n and ΔT. They are used to compute the evolution step of the EnKF. Since no reactive power consumption dataset was available, a similar model is assumed for reactive power changes. However, active and reactive power consumption changes are assumed to be independent, which is a common assumption in DSSE literature. The proposed method is generic and can be applied to any dataset from across the globe.

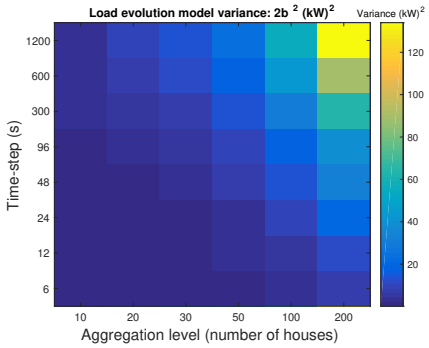


Fig. 4. Influence of aggregation level and time step on the scale parameter b. (The figure is best viewed on screen)

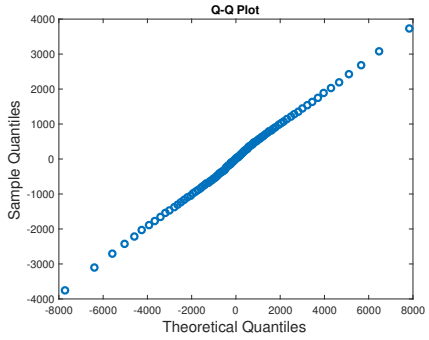


Fig. 5. Q-Q plot illustrating a good fit of the Laplace distribution ($n = 200$, $\Delta T = 6$s.).

B. Test Distribution System

The 33-bus test feeder data includes active and reactive power loads at each bus; bus-1 is the substation transformer bus, with V_0 set to 12.66 kV. The number of houses n_i aggregated at a bus i is selected such that $n_i = n_{11} P_i^{33bus} / P_{11}^{33bus}$

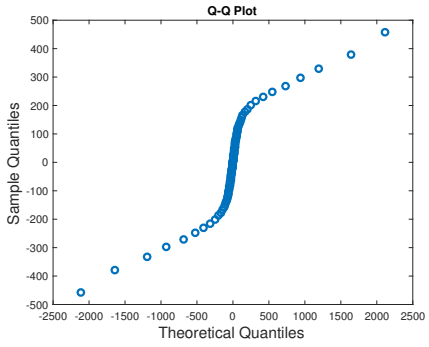


Fig. 6. Q-Q plot illustrating the worst fit ($n = 10$, $\Delta T = 6$s.) of all the fits considered

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

where $n_{11} = 10$ houses and P_i^{33bus} is the static 33-bus active power load at bus i.

The corresponding distribution transformer traces are generated from the second half of the dataset, by summing the desired number of profiles, picked randomly. Each trace is then scaled so that the mean of the profile matches the load values. The values given by the empirical function in Section V-A are scaled accordingly. Because no dataset for reactive power consumption is available, active and reactive power profiles are generated independently from the same dataset.

C. Measurement Model

The simulation models used for measurements are described in this section.

PMU: the PMU measurement error is simulated as an additive white Gaussian noise of nominal variance σ^2_{PMU}, for both voltage magnitudes and angles. The readings $\tilde{V}_s$ and $\tilde{\delta}_s$ provided by the PMU at each bus s ($s \in S \subseteq I$) have an error variance such that $\mathbf{E}[\tilde{a}^2] = \sigma^2_{PMU} \cdot \tilde{a}^2$, where $\tilde{a}$ indicates either the voltage magnitude or angle. The measurement errors are independent across buses, and the voltage magnitude error independent of the angle error. The PMU resolution is set to 1% ($\sigma_{PMU} = 0.01$); the PMU placement map $\mathscr{S}$ is determined using a greedy method [8], i.e., PMUs are sequentially added at the location that provides the most improvement (with 32 load buses, a maximum of 32 PMUs). The placement of PMUs is beyond the scope of this work; many researchers have addressed this issue, see for example [32]. The sequential bus placement map used is the following: $\mathscr{S} = \{33,\ 32,\ 31,\ 18,\ 17,\ 30\ 16,\ 29,\ 15,\ 14,\ 13,\ 28,\ 12,\ 11,\ 10,\ 9,\ 8,\ 27,\ 26,\ 7,\ 6,\ 25,\ 24,\ 5,\ 4,\ 23,\ 3\ 22,\ 21,\ 20,\ 19,\ 2\}$

Pseudo-measurements: the forecasts P_i^f and Q_i^f are taken as the mean value of the load profile generated at each distribution transformer i, as in [8]. They are constant over the simulated period. Using the training set, the nominal standard deviation of the forecast was evaluated and set to $\sigma_0 = 30\%$, for both active and reactive powers, irrespective of the aggregation level. Therefore for each bus i, $\sigma_i^{fp} = \sigma_0 P_i^f$ and $\sigma_i^{fq} = \sigma_0 Q_i^f$ (6). The constant apparent power forecast $|S_i^f|$ is such that $|S_i^f| = |P_i^f + jQ_i^f|$. Finally each σ_i^f (18) is computed as $\sigma_i^f = \sigma_0 |S_i^f|$. Pseudo-measurements to which synthetic perturbations following a Gaussian distribution are used as "best-guess" initial ensemble.

Error time-correlation: ψ_i^p and ψ_i^q are evaluated as follows: since the same data is used for generating the active and reactive power profiles, ψ_i^p and ψ_i^q are equal. They are evaluated on the training set. Given an aggregation level and a time-step length, load profiles are built. The autocorrelation function R_i^e of the difference between the profile and its mean (representing the forecast error) is computed. The value of ψ_i^p and ψ_i^q is given by $R_i^e(\Delta T)$.

D. Validation

The theoretical and simulation results are presented in Figs. 7a-7c, obtained by averaging the results of several realizations. A realization is defined as the observed performance of both the WLS and PASE on the 33-bus system. For each realization, new load profiles are generated based on the testing set, while the other parameters stay the same. The performance of the WLS and PASE are plotted alongside with the theoretical ones in Fig. 7a, where a time-step of 6 seconds has been used. WLS has been studied in [8] using synthetic data. Similar trends are observed here with real data. Note that since WLS is snapshot-based, the size of the time-step does not matter. For PASE, the theoretical results are close to the actual performance observed in simulation as the number of PMUs introduced in the system increases, which validates the theoretical approach. Similar trends are observed for different time-steps. The actual gain brought about by PASE is compared with the theoretical one in Fig. 7b for a time step of 6 seconds. The theory allows to estimate quickly the order of magnitude of the gain achieved by PASE over WLS. Finally the influence of the time-step on the gain is compared in Fig. 7c for two PMU configurations (5 PMUs and 20 PMUs). Theory and simulation follow the same trend. The gap between theory and simulation is relative to that observed in Fig. 7b.

Clearly, in the theoretical formulation, the linearization process and the simple load modeling introduce a visible error, especially when the number of PMUs is small. However, such error is tolerated as the goal of the theoretical computations is to provide a rough estimate of the gain from using PASE over WLS, without having to carry out expensive Monte-Carlo simulations.

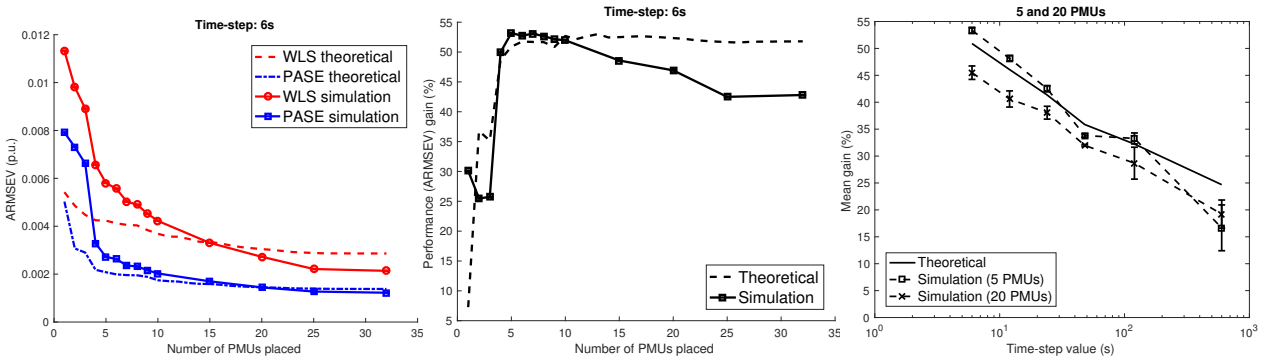
E. Comparison Between WLS and Proposed PASE Method

The results presented in Fig. 7a illustrate the improvements achieved by the proposed PASE method. Clearly, using a load evolution model improves the performance of the estimator; given an arbitrary target error of 0.004 p.u., WLS requires more than 10 PMUs while PASE only 4. Even when each bus of the distribution system is monitored by a PMU, the proposed PASE method still brings about an improvement of more than 40% when using a time-step of 6 seconds. As illustrated in Fig. 7c, higher gains are obtained for smaller time-steps. Indeed, for larger time-steps, the load has more chances of changing by a large magnitude between two estimates and thus has less inertia. Even for large time-step (e.g., 10 mins) there is a gain of about 15%. In practice, the granularity of the time-step depends on the available computational speed. The smallest time-step considered in this work is 6 seconds and represents a lower-bound on what was tried out. In comparison, the DSSE problem was solved at each step in under 200 millisecond. Clearly, PASE has a higher computational cost than the SoA. Indeed L power-flows must be solved, and the cost of solving one of those power-flow increases with the number of buses in the system. The reported performance above is for $L = 500$ and 33 buses, where all the computations are performed using Matlab on a standard desktop computer. If the computational cost becomes a limiting factor, diverse strategies exist such as running the power-flow computations in parallel (the L power-flows are independent).

F. Engineering Insights

In practice, the LDC will need to make trade-offs in the choice of the following parameters: number of PMUs, their

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems



(a) (b) (c)

Fig. 7. (a) ARMSEV value function of the number of PMUs. The performance of the proposed PASE method is compared with WLS. The theoretical results are also compared against simulation results. A lower value means better performance. (b) Comparison of the gain from using PASE over WLS on ARMSEV depending on the number of PMUs. The theoretical results are compared to the observed gain in simulation. (c) Influence of the time-step on the mean performance gain. The theoretical results are compared to the observed gain in simulation. The time-step axis has a logarithmic scale. The error bars represent the variance.

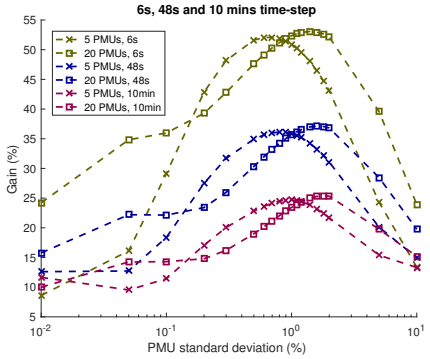


Fig. 8. Influence of PMU quality on the (theoretical) gain achieved by PASE over WLS. The PMU accuracy is characterized by its measurement error standard deviation, a lower value means a more accurate PMU. A logarithmic scale is used for the variance axis.

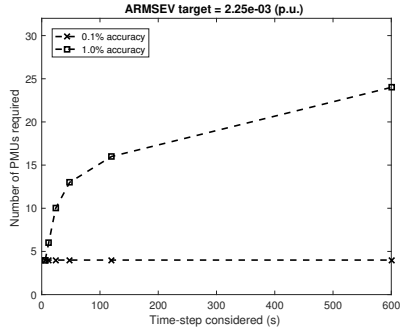


Fig. 9. Minimum number of PMUs required to achieve an average target error of $2.25e-3$ p.u., function of time-step and for two PMU accuracies (computed using the theoretical formulation).

accuracy and the time scale. The influence of PMU accuracy on the theoretical gain achieved by PASE is shown in Fig. 8, the three parameters considered are depicted in the plot. The maximum gain is attained for a PMU error variance of about 1%. Clearly as the PMU measurement standard deviation decreases (i.e., the PMU becomes more and more accurate) the gain achieved by PASE decreases since the load evolution model is not as useful in such circumstances. Similarly, as the standard deviation of the PMU increases, the gain decreases, since the load evolution model has to compensate for both poor forecast accuracy and poor PMU measurement accuracy. This figure also illustrates the role of the time-step, the gain achieved by the filtering technique decreasing as the time-step increases, underlining the limits of the load evolution model.

Note that one should be careful not to draw any relation or trend between the number of PMUs and the gain achieved in Fig.. 7b. As illustrated by Fig. 8 the gain in that case would also depends on the PMU accuracy.

The trade-off between the three parameters considered is illustrated by Fig. 9: two PMU accuracies are used to draw the plots. An arbitrary target error is fixed and the minimum number of PMUs required is determined as a function of the time-step. Clearly, the time-step has little influence on a very accurate PMU. However, the more accurate the PMU, the more costly it will be. With the same number of PMUs placed in the system (4), choosing a PMU ten times less accurate will provide the same performance given that a time-step small enough (6 seconds) is chosen.

G. Sudden Load Changes and Bad Data Detection

Sudden load changes, topology errors and gross bad data in measurements constitute irregularities that are susceptible to affect state estimators that are not robust enough. Some estimators have been specifically designed to be insensitive to these issues, such as the one presented in [33].

The robustness of PASE is studied in the following case: an artificial sudden load change is simulated at bus 12. Without

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2018.2847289, IEEE Transactions on Power Systems

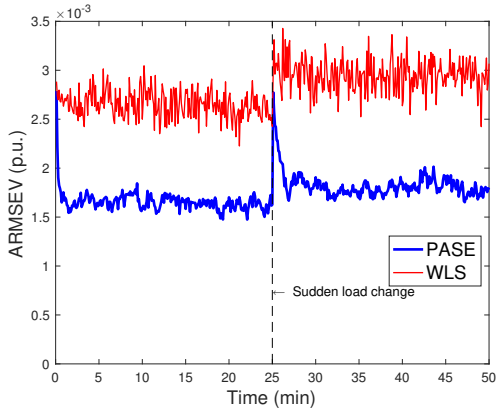


Fig. 10. Impact of sudden load change at bus 12 on the ARMSEV value for PASE (bottom blue curve) and WLS (top red curve). The simulation parameters are as follow: 10 PMUs, $\delta T = 6$s, 1% PMU error standard deviation.

any warning, the active and reactive power consumptions are divided by a factor 10, to mimic a sudden and unexpected load drop. As one would expect, the pseudo-measurements cannot foresee such event and their values are kept as if the system was under normal operations. 10 PMUs are placed in the system, and a time-step of 6 second is used. ARMSEV is used to assess the performance. For comparison purposes, the performance of WLS is displayed as well. The goal of such comparison is not to claim that PASE is more robust than WLS, but rather show with a single experiment that PASE has enough intrinsic robustness to handle sudden load changes. Its robustness will be studied in more details in a future work. The results, averaged over several realizations, are presented in Fig. 10. During the first 25 minutes, where the system is under normal operating conditions, the ARMSEV achieved by PASE is about 40% lower than the one achieved by WLS. When sudden load drop happens (at t = 25 mins), the average error of PASE surges, until the filter tracks again correctly the new operating point. Regarding WLS, the average error of the estimator increases at the time of the incident and remains higher than before.

To ensure estimation consistency, topological errors and bad data (erroneous measurements) must be detected and either corrected or removed. For WLS, bad data detection is usually performed by doing hypothesis testing on residuals [2] (i.e., the objective function is compared to a maximum threshold value). In the proposed PASE framework, typical bad data detection can be performed as follows: first WLS is run. The estimated state is discarded and only bad data detection is performed. Once input data is clean, the system state is computed using PASE. Such bad data detection mechanism is simple. Other more advanced bad data detection methods in filtered framework have also been developed, such as the one described in [34], where the authors describe the use of the state forecasting capability of Kalman filters for improved bad data detection.

VI. Conclusions

A novel PASE method for DSSE and its analysis framework were presented. The PASE method performs the fusion of measurements and pseudo-measurements and requires fewer PMUs than WLS to achieve the same estimation error, for time-steps under 15 minutes. Engineering insights were presented highlighting the major trade-offs in the choice of decision variables for the LDC. Using a smaller time-step allows the LDC to relax the requirements on the PMU quality and their number. There are several remaining challenges, such as a further study of the state forecasting capabilities of PASE for bad data detection, the study of the influence of distributed generation and its modeling as well as the impact of an unbalanced system on PASE. It would also be interesting to assess the performance of PASE on meshed systems.

References

[1] D. Atanackovic and V. Dabic, "Deployment of real-time state estimator and load flow in BC Hydro DMS - challenges and opportunities," in *2013 IEEE Power Energy Society General Meeting*, 2013.
[2] A. Monticelli, *State Estimation in Electric Power Systems*. Springer, 1999.
[3] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Micro-synchrophasors for distribution systems," in *ISGT, IEEE PES*, 2014.
[4] R. N. Rodrigues, J. K. Zatta, P. C. C. Vieira, and L. C. M. Schlichting, "A low cost prototype of a Phasor Measurement Unit using Digital Signal Processor," in *IEEE Biennial Congress of Argentina*, Jun. 2016.
[5] C. A. Fantin, M. R. C. Castillo, B. E. B. d. Carvalho, and J. B. A. London, "Using pseudo and virtual measurements in distribution system state estimation," in *IEEE PES T&D-LA*, 2014.
[6] A. Primadianto and C. N. Lu, "A Review on Distribution System State Estimation," *IEEE Transactions on Power Systems*, 2016.
[7] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, "Distribution circuit state estimation using a probabilistic approach," *IEEE Transactions on Power Systems*, 1997.
[8] L. Schenato, G. Barchi, D. Macii, R. Arghandeh, K. Poolla, and A. V. Meier, "Bayesian linear state estimation using smart meters and PMUs measurements in distribution grids," in *SmartGridComm*, 2014.
[9] S. Alam, B. Natarajan, and A. Pahwa, "Distribution Grid State Estimation from Compressed Measurements," *IEEE Trans. Smart Grid*, 2014.
[10] H. Wang and N. N. Schulz, "A revised branch current-based distribution system state estimation algorithm and meter placement impact," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 207–213, Feb. 2004.
[11] C. Klauber and H. Zhu, "Distribution system state estimation using semidefinite programming," in *NAPS*, 2015.
[12] M. B. D. C. Filho and J. C. S. d. Souza, "Forecasting-Aided State Estimation - Part I: Panorama," *IEEE Trans. on Power Systems*, 2009.
[13] S.-C. Huang, C.-N. Lu, and Y.-L. Lo, "Evaluation of AMI and SCADA Data Synergy for Distribution Feeder Modeling," *IEEE Trans. Smart Grid*, 2015.
[14] S. Sarri, M. Paolone, R. Cherkaoui, A. Borghetti, F. Napolitano, and C. A. Nucci, "State estimation of active distribution networks: comparison between WLS and iterated Kalman-filter algorithm integrating PMUs," in *Innovative Smart Grid Technologies (ISGT Europe)*, 2012.
[15] R. Singh, B. C. Pal, and R. A. Jabr, "Choice of estimator for distribution system state estimation," *IET Gener. Transm. Distrib.*, 2009.
[16] G. Evensen, "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, Nov. 2003.
[17] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. on Power Del.*, 1989.
[18] S. Lefebvre, J. Prevost, J. C. Rizzi, P. Ye, B. Lambert, and H. Horisberger, "Operational experience with state estimation at hydro-quebec," 2008.
[19] Y. Chen, M. J. Rice, K. R. Glaesemann, S. Wang, and Z. Huang, "Sub-second parallel state estimation," Pacific Northwest National Lab.(PNNL), Richland, WA (United States), Tech. Rep., 2014.

[20] I. Shah and F. Lisi, "Day-ahead electricity demand forecasting with non-parametric functional models," in *2015 12th International Conference on the European Energy Market (EEM)*, 2015.
[21] Y. Besanger, R. Caire, O. Chilard, and P. Deschamps, "Distribution state estimation performances with different state vectors," in *PowerTech (POWERTECH), 2013 IEEE Grenoble*, 2013.
[22] A. Abur and A. G. Expsito, *Power System State Estimation: Theory and Implementation*. CRC Press, Mar. 2004.
[23] B. Hayes and M. Prodanovic, "State estimation techniques for electric power distribution systems," in *2014 European Modelling Symposium*, 2014.
[24] J. W. Taylor and P. E. Mcsharry, *Short-Term Load Forecasting Methods: An Evaluation Based on European Data*.
[25] J. W. Taylor, L. M. de Menezes, and P. E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *International Journal of Forecasting*, 2006. [Online]. Available: http://eprints.maths.ox.ac.uk/281/
[26] J. Hinman and E. Hickey, "Modeling and forecasting short-term electricity load using regression analysis," *Journal of IInstitute for Regulatory Policy Studies*, 2009.
[27] K. Wang, Y. Li, and C. Rizos, "Practical Approaches to Kalman Filtering with Time-Correlated Measurement Errors," *IEEE Trans. Aerosp. Electron. Syst.*, 2012.
[28] M. G. Petovello, K. OKeefe, G. Lachapelle, and M. E. Cannon, "Consideration of time-correlated errors in a Kalman filter applicable to GNSS," *Journal of Geodesy*, 2009.
[29] B. Anderson and J. Moore, *Optimal Filtering*. Prentice-Hall, 1979.
[30] J.-H. Teng, "A direct approach for distribution system load flow solutions," *IEEE Transactions on Power Delivery*, Jul. 2003.
[31] O. Ardakanian, S. Keshav, and C. Rosenberg, "Markovian Models for Home Electricity Consumption," in *ACM SIGCOMM*, 2011.
[32] R. Singh, B. C. Pal, and R. B. Vinter, "Measurement Placement in Distribution System State Estimation," *IEEE Trans. on Power Syst.*, 2009.
[33] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended kalman filter for power system dynamic state estimation," *IEEE Transactions on Power Systems*, vol. 32, 2017.
[34] G. Valverde and V. Terzija, "Unscented kalman filter for power system dynamic state estimation," *Transmission Distribution IET Generation*, vol. 5, no. 1, pp. 29–37, Jan. 2011.

Côme Carquex received a Master's degree in Electrical and Computer Engineering at the University of Waterloo, Waterloo, ON, Canada, in 2017.

Catherine Rosenberg, FIEEE is a Professor with the Department of Electrical and Computer Engineering at the University of Waterloo and the Canada Research Chair in the Future Internet. Since April 2018, she is also the Cisco Research Chair in 5G Systems. She was elected a Fellow of the Canadian Academy of Engineering in 2013. Her research interests are in networking, wireless, and energy systems. More information is available at https://ece.uwaterloo.ca/˜cath/.

Kankar Bhattacharya (M'95-SM'01–F'17) received the Ph.D. degree in electrical engineering from the Indian Institute of Technology, New Delhi, India, in 1993. He was in the faculty of Indira Gandhi Institute of Development Research, Mumbai, India, during 19931998, and then the Department of Electric Power Engineering, Chalmers University of Technology, Gothenburg, Sweden, during 19982002. He joined the E&CE Department of the University of Waterloo, Waterloo, ON, Canada, in 2003 where he is currently a Full Professor. His research interests are in power system economics and operational aspects.