Statistical Error Analysis of Household Load Profile for Medium Voltage Grid State Estimation

Yu Xiang, J.F.G. Cobben, P.F. Ribeiro
Faculty of Electrical Engineering
Eindhoven University of Technology
Eindhoven, Netherlands
y.xiang@tue.nl

Abstract—The paper comprises a detailed analysis of the error ranges of two commonly proposed load estimation methods for household loads, based on statistics. Firstly, a household load profile model is built taking individual user preference into account. To analyze the error range, it is divided into system error (representing proportion factors) and random error (representing the diversity of stochastic loads and measurement inaccuracies). For both estimation methods, two kinds of bounds are used to evaluate the error range, with central limit theorem. Finally, comparisons between different errors are made using field data from a Dutch utility.

Index Terms—distribution grid state estimation, household load profile, load estimation, central limit theorem

I. INTRODUCTION

A distribution management system (DMS) is required for a modern medium voltage (MV) distribution grid. The heart of this system is a state estimator, which evaluates the state of the grid based on input data. However, currently, the measurement data in MV grids are much less than redundant for the state estimator to achieve observability. Thus more measurement devices should be installed. Besides that, many studies suggested using load prediction (estimation) as an additional pseudo measurement for the state estimator [1].

The load estimation is especially suitable for household and commercial loads. Several load estimation methods have been proposed and verified with empirical data [1]–[3]. But besides the estimated values of loads, their error range is also essential for the state estimator to determine the weighting factor in the optimization algorithm. To solve this problem, [4] adopted an artificial neural network approach for the pseudo measurement modeling.

To provide a simple and robust solution for the grid operators, this paper analyzes the error ranges of two commonly proposed load estimation methods for household loads, based on statistics. Section II defines the problem and sets the research questions. A modeling of household loads is presented in Section III. Section IV presents the load estimation methods and analyzes their error ranges, using central limit theorem. To demonstrate the feasibility of error analysis, a case study with field data from a Dutch utility is given in Section V.

II. PROBLEM DEFINITION

A sample MV grid is shown in Fig. 1. Along several feeders from an MV bus there are many MV points of connections (PoCs), which are connected with household (H), commercial (C), or industrial (I) loads. In general, it is suggested that all the industrial loads should be directly measured, due to their complex behavior and significant impact on the grid. Therefore, the load estimation is more relevant for household and commercial loads. The paper discusses household loads, although commercial loads can be analyzed with a similar approach.

The real-time active loads are measured at a limited number of household PoCs, and the active loads at other household PoCs should be estimated based on real measurement data and stochastic load characteristics. As shown in Fig. 2, the estimated power flow is considered as a pseudo measurement for the state estimator.

It has been suggested that the loads at the unmeasured PoCs can be calculated in proportion with the loads at the measured PoCs. To obtain the proportion factors, four methods have been proposed [1]–[3], based on:

1) number of customers
2) annual bill
3) historical maximum load
4) rating of the distribution transformer

These methods have been verified with empirical results. Generally methods 1 and 2 are more accurate and practical [3]. However, load estimation should provide not only the values of real-time loads, but also their error range, which is very important for the state estimator. Thus, the paper comprises a statistical analysis of the error ranges of estimation methods 1 and 2 for household loads. The following research questions are essential:

1) What does the system and random error of load estimation represent respectively?
2) Considering system and random error, what approach is suitable to evaluate the estimation error?
3) Can the estimation error be reduced significantly by increasing the number of PoCs with real measurements?
4) Does the placement of real measurements influence the accuracy of load estimation?

III. HOUSEHOLD LOAD MODELING AND MEASUREMENT

In order to analyze the problem, the stochastic characteristics of household loads should be known first. In this section,
a model of household load profile is presented with several important parameters. In addition, a brief analysis is given for the error and measuring range of load measurements.

A. Household Load Profile

Let \( H \) be the set of all PoC numbers that is connected with household loads. For any PoC \( i \in H \), there are \( n_i \) LV household customers in total, as shown in Fig. 1. \( P_{i,j}(t) \) is defined as the active power curve of the \( j^{th} \) customer in MV PoC \( i \). It is assumed that each \( P_{i,j}(t) \) has the following form:

\[
P_{i,j}(t) = \xi_{i,j}X_{i,j}(t)
\]

where \( \xi_{i,j} \) is the user preference factor, and \( X_{i,j}(t) \) is the base load profile.

The user preference factor \( \xi_{i,j} \) represents the individual characteristic of electricity usage per customer; some customers naturally use more electricity and others use less. The mean (mathematical expectation) of all the user preference factors should be 1, shown in (2). And \( \lambda^2 \) is defined as the variance of user preference factor. Statistical smart metering data [5] show that the value of \( \lambda \) is between 0.2–0.6.

\[
E\{\xi_{i,j}\} = 1, \quad D\{\xi_{i,j}\} = \lambda^2
\]  

Fig. 1. Sample grid with a number of MV PoCs, including household (H), commercial (C), and industrial (I) loads.

Fig. 2. Input of the state estimator, including network parameters, real and pseudo measurements.

The square of user preference factor \( \xi_{i,j}^2 \) plays an important role. The mean \( \phi_d \) and variance \( \lambda_d^2 \) are defined below.

\[
\phi_d = E\{\xi_{i,j}^2\}, \quad \lambda_d^2 = D\{\xi_{i,j}^2\}
\]  

The base load profile \( X_{i,j}(t) \) is a stochastic process which represents the common human behavior of the electricity usage. Though the actual values of every \( X_{i,j}(t) \) may be different at any given time point, their probability distributions are identical and independent with each other.

The mean function \( P_m(t) \), calculated in (4), represents the average electricity consumption pattern. Fig. 3 shows an example of \( P_m(t) \) during the winter and summer period in Netherlands.

\[
P_m(t) = E\{X_{i,j}(t)\}
\]  

The covariance function \( K(t_1, t_2) \) is the covariance between the value of \( X_{i,j}(t) \) at two time points \( t_1 \) and \( t_2 \). To achieve a mathematical assumption, [6]–[8] suggest that the covariance function has the “Squared Exponential” form:

\[
K(t_1, t_2) = \sigma(t_1)\sigma(t_2)e^{-\frac{(t_1-t_2)^2}{2\sigma^2}}
\]  

The time constant \( T \) represents: “Within how much time the user electricity usage is correlated.” Particularly, when \( t_1 = t_2 \), the \( K(t_1, t_1) = \sigma^2(t_1) \) is the self-variance at time point \( t_1 \), and \( \sigma(t_1) \) is the standard deviation.

The following values are defined for further analysis. Annual mean and maximum of \( P_m(t) \):

\[
\mu_m = \frac{1}{\text{year}} \int_\text{year} P_m(t)dt, \quad \mu_{\text{max}} = \max_{t \in \text{year}} \{P_m(t)\}
\]  

Annual maximum standard deviation of \( X_{i,j}(t) \):

\[
\sigma_{\text{max}} = \max_{t \in \text{year}} \{\sigma(t)\}
\]  

Average power and standard deviation at time point \( t_0 \):

\[
\mu_0 = P_m(t_0), \quad \sigma_0 = \sqrt{K(t_0, t_0)} = \sigma(t_0)
\]  

Ignoring the loss in LV grid, the total real-time active power at PoC \( i \in H \), at any given time point \( t_0 \):

\[
P_i = \sum_{j=1}^{\infty} P_{i,j}(t_0)
\]
B. Load Proportion Factor

For any $i \in H$, the load proportion factor $\xi_i$ is defined as the sum of all the user preference factors at PoC $i$:

$$\xi_i = \sum_{j=1}^{n_i} \xi_{i,j}$$  \hspace{1cm} (10)

Since every $P_{i,j}$ in the same PoC $i \in H$ is independent of each other. The mean and variance of the total active power $P_i$ at time point $t_0$ can be calculated by (11) and (12), where $\xi_{i,d}^2$ is defined as the square proportion factor.

$$E\{P_i\} = \xi_i \mu_0$$  \hspace{1cm} (11)

$$D\{P_i\} = \xi_{i,d}^2 \sigma_0^2 \quad \xi_{i,d}^2 = \sum_{j=1}^{n_i} \xi_{i,j}^2$$  \hspace{1cm} (12)

The mean and variance of $\xi_i$ and $\xi_{i,d}^2$ can be calculated by (13) and (14) respectively, based on (2) and (3).

$$E\{\xi_i\} = n_i \quad D\{\xi_i\} = n_i \lambda^2$$  \hspace{1cm} (13)

$$E\{\xi_{i,d}^2\} = n_i \phi_d \quad D\{\xi_{i,d}^2\} = n_i \lambda_d^2$$  \hspace{1cm} (14)

C. Estimated Relative Load Proportion Factor

According to (11), for any $i, j \in H$, the mathematical expectations of the real-time active power $P_i$ and $P_j$ have the following relation:

$$E\{P_i\} = \xi_i \quad \xi_i = \xi_{i,d}^2$$  \hspace{1cm} (15)

The relation in (15) is the fundamental basis of the load estimation approach in this paper. It allows the grid operator to measure the real-time power at some of the household PoCs, and estimate the real-time power at other household PoCs based on the load proportion factor. For example, if meters are installed at PoC $x \in H$, for the active power at another PoC $i$, (15) can be reformed to (16), where $\hat{\theta}_i$ is the relative load proportion factor.

$$E\{P_i\} = \hat{\theta}_i E\{P_x\} \quad \hat{\theta}_i = \frac{\xi_i}{\xi_x}$$  \hspace{1cm} (16)

Unfortunately, the exact user preference factors and load proportion factors are unknown. They can however be estimated based on other data. Several papers summarized and verified four estimation methods [1]–[3], while the top two methods are discussed in this paper: estimation based on the number of customers and estimation based on the annual bill. The estimated relative load proportion factor $\hat{\theta}_i$ is defined using these two methods.

1) Method 1: estimation based on number of customers. Let PoC $x \in H$ have the most customers, $\hat{\theta}_i$ is defined as:

$$\hat{\theta}_i = \frac{n_x}{n_x}$$  \hspace{1cm} (17)

2) Method 2: estimation based on annual bill. The paper defines $C_i$ as the total annual consumption of all customers at PoC $i$. Similarly, let PoC $x \in H$ have the largest total annual consumption, $\hat{\theta}_i$ is defined as:

$$\hat{\theta}_i = \frac{C_i}{C_x}$$  \hspace{1cm} (18)

According to the definitions above, no matter which estimation method is used, the following statement is always true:

$$\hat{\theta}_i \leq \hat{\theta}_x = 1$$  \hspace{1cm} (19)

D. Load Measurement and Error

Along all the PoCs, let $S \subset H$ be the set of household PoC numbers where their active power is measured. Depending on the estimation method for the load proportion factor, it is important to install measurement devices at least at the PoC $x$, or $x \in S$.

The active power measurement error at PoC $r \in S$ is noted as $\zeta_r$, as shown in (20). Usually $\zeta_r$ is assumed to be approximately normally distributed [9]. Given that the system error is corrected by modern measurement devices, the mean of $\zeta_r$ is zero, and the standard deviation $\delta_r$ of the error can be calculated in (21), where $\Psi_r$ is the measuring range and $\tau_0$ is the relative error. Almost all ($99.7\%$) the $\zeta_r$ falls in $\pm 3\delta_r$.

$$\tilde{P}_r = P_r + \zeta_r$$ \hspace{1cm} (20)

where $P_r$ is the real value of the active power, and $\tilde{P}_r$ is the measured value.

$$\zeta_r \sim \mathcal{N}(0, \delta_r^2) \quad \delta_r = \sqrt{D\{\zeta_r\}} = \frac{1}{3} \Psi_r \tau_0$$ \hspace{1cm} (21)

Obviously, the absolute measurement error can be reduced by decreasing the measuring range, which can be done in any modern digital measurement device. Thus, the paper proposes limit the measuring range to cover the largest measured value with sufficient margin. Nevertheless, the measuring range is usually adjusted discretely. In the paper the measuring range of active power at PoC $r \in S$ is set to be two times its maximum active power, as shown in (22).

$$\Psi_r = 2P_{r,\text{max}} \approx 2n_r \mu_{\text{max}}$$ \hspace{1cm} (22)

IV. LOAD ESTIMATION AND ERROR ANALYSIS

A. Active Power Estimation

The active power at an unmeasured PoC is estimated in proportion with the active power at the measured PoCs, based on the relative load proportion factor $\hat{\theta}_i$. For any $i \in H \setminus S$, the real-time active power $\tilde{P}_i$ is estimated by (23).

$$\tilde{P}_i = \Theta_i \sum_{r \in S} \tilde{P}_r, \quad \Theta_i = \frac{\hat{\theta}_i}{\sum_{r \in S} \hat{\theta}_r}$$ \hspace{1cm} (23)
B. Random Error and System Error

The estimation error $\alpha_i$ is defined as the deviation between $\hat{P}_i$ and $P_i$, as shown in (24). The mean of $\alpha_i$ represents the system error, and according to (11) and (20), it can be calculated by (25), where $\beta_i$ is the system error factor. Additionally, the variance of $\alpha_i$ represents the random error. With (12) and (21), it can be calculated by (26).

$$\alpha_i = \hat{P}_i - P_i \quad (24)$$

$$E\{\alpha_i\} = \beta_i \mu_0, \quad \beta_i = \Theta_i \sum_{r \in S} \xi_r - \xi_i \quad (25)$$

$$D\{\alpha_i\} = \Theta_i^2 \sum_{r \in S} (\xi_{i,r}^2 \sigma_0^2 + \delta_r^2) + \xi_{i,r}^2 \sigma_0^2 \quad (26)$$

Using the formulas above, research question 1 can be answered: the system error is generated when estimating the load proportion factors, while the random error is caused by the diversity of stochastic loads and measurement inaccuracies.

Additionally, when more PoCs are measured, the first item in (26) remains the same. Thus, this answers research question 3: increasing the number of measured PoCs can decrease the estimation error, but this is limited by the diversity of the random loads at PoC $i$ itself. In Section V the paper shows when this limit is reached.

Moreover, (19) and (23) show that $\Theta_i$ is always smaller than 1. Therefore, the random error of estimation in (26) is limited by $\Theta_i$. This answers research question 4: at a minimum the PoC with the largest $\hat{\theta}_i$ should be measured, to ensure the robustness of the estimation algorithm.

C. Evaluating Error using Central Limit Theorem

Analyzing $\alpha_i$ is challenging because the probability distribution of $X_{i,j}(t)$ is unknown. However, usually each PoC contains many LV household customers, which means each $P_i$ is the sum of a large number of independent random variables. Therefore, according to central limit theorem, it can be assumed that each $P_i$ follows normal distribution. As a result, the estimation error $\alpha_i$ is also normally distributed. Thus, an upper bound of $\alpha_i$ at a confidence level of 0.997 (6-sigma bound) can be calculated by (27).

$$|\alpha_i| \leq |\beta_i|_{\max} \mu_{\max} + 3 \sqrt{D\{\alpha_i\}} \quad (27)$$

In calculating the mean and variance of $\alpha_i$, the distribution of $\xi_i$ and $\xi_{i,d}^2$ are still unknown, but $\xi_i$ and $\xi_{i,d}^2$ are also assumed to be normally distributed. Central limit theorem can again be applied to estimate the maximum value of $\xi_i$ and $\xi_{i,d}^2$. With (14), (21) and (26), an upper bound of $D\{\alpha_i\}$ can be found at a confidence level of 0.997, as shown in (28).

$$D\{\alpha_i\} \leq \sigma_{\max}^2 \Theta_i^2 \left( \phi_d N_S + 3 \lambda_d \sqrt{N_S} \right)$$

$$+ \left( \frac{2}{3} \mu_{\max} \theta_0 \right)^2 \Theta_i^2 N_{S,d}^2$$

$$+ \sigma_{\max}^2 \left( \phi_d n_i + 3 \lambda_d \sqrt{n_i} \right) \quad (28)$$

where:

$$N_S = \sum_{r \in S} n_r, \quad N_{S,d}^2 = \sum_{r \in S} n_r^2 \quad (29)$$

Similarly, with (13), (17) and (25), an upper bound of system error factor $\beta_i$ for Method 1 can be obtained:

$$|\beta_i| \leq 3 \lambda \sqrt{\frac{n_i^2}{N_S}} + n_i \quad (30)$$

The estimation of $\beta_i$ for Method 2 is more complex, because $\hat{\theta}_i$ and $\xi_i$ are correlated in (18). This relation should be considered in the analysis. The paper omits the detail of this trivial mathematical derivation. As a result, an upper bound of $\beta_i$ for Method 2 can be obtained by (31).

$$|\beta_i| \leq \sqrt{\phi_d \left( \Theta_i^2 N_S + n_i \right)} + 3 \lambda_d \left( \Theta_i^2 \sqrt{N_S} + \sqrt{n_i} \right) \frac{3 \sigma_{\max} \sqrt{\Gamma_i \mu_{\max}}}{\mu_{\max}} \quad (31)$$

$$\Gamma_i = \int_{-1}^{1} \int_{-1}^{1} e^{-\left( \frac{t_1-t_2}{\sqrt{2}} \right)^2} dt_1 dt_2 \quad (32)$$

With (28), (30) and (31), an upper bound of $\alpha_i$ can be evaluated by (27). However, central limit theorem is applied twice in the approach: the maximum value of the mean and variance of $\alpha_i$ are separately obtained first, to calculate the maximum value of $\alpha_i$. Thus, this approach is called a conservative bound in this paper.

D. Evaluating Error with Progressive Bound

The conservative bound considers system and random error separately to ensure its validity of the error evaluation, but this might be too conservative sometimes, leading to very large margin over the actual error range. Thus, the effect of pseudo measurement data could be hindered in the state estimation. To solve the problem, a progressive bound is proposed in this paper. Compared to the conservative bound, the progressive bound has the following modifications:

- The system and random error are considered together as the sum of independent random variables. As a result, the mean of $\alpha_i$ is supposed to be zero, and the variance consists of a system error part and a random error part. Central limit theorem is applied only once, and the maximum value of $\alpha_i$ (at confidence level of 0.997) is evaluated as three times standard deviation.
- The average value (instead of maximum value) of $\xi_{i,d}^2$ is used when calculating the variance, to avoid double maximization.

With the modifications above, a progressive upper bound of $\alpha_i$ can be calculated by (33). The variance of $\alpha_i$ consists of three parts. $D_{RL_i}$ represents the random error from the diversity of random loads, evaluated by (34). $D_{R2,i}$ is the random error from measurement inaccuracies, evaluated by
According to smart metering data, it can be assumed that each function of the user preference factor is shown in Fig. 4. Over the robustness, the case study in Section V shows the therefore, it is suitable when the performance is preferred but is at risk of being lower than the actual error range. Thus, it is suitable for situations when the robustness of state estimation error with a large margin, but ensures its validity.

\[ |\alpha_i| \leq 3 \sqrt{D_{R1,i} + D_{R2,i} + D_{S,i}} \]  
(33)

where:

\[ D_{R1,i} \approx \sigma_{\text{max}}^2 \phi_d(\Theta_i^2 N_S + n_i) \]  
(34)

\[ D_{R2,i} \approx \frac{4}{9} \sigma_{\text{max}}^2 \phi_d(\Theta_i^2 N^2_{S,d}) \]  
(35)

Method 1:

\[ D_{S,i} \approx \mu_{\text{max}}^2 \lambda^2 \left( \frac{n_i^2}{N_S} + n_i \right) \]  
(36)

Method 2:

\[ D_{S,i} \approx \mu_{\text{max}}^2 \sigma_{\text{max}}^2 \Gamma_2 \phi_d(\Theta_i^2 N_S + n_i) \]  
(37)

Regarding to two kinds of bounds, research question 2 can be answered: the conservative bound provides a boundary close to the actual error range, but is at risk of being lower than the actual error range. Therefore, it is suitable when the performance is preferred over the robustness. The case study in Section V shows the difference between these two bounds.

V. CASE STUDY WITH TYPICAL PARAMETERS

In this section a case study is given with typical parameters of a Dutch household load profile. The real error range of load estimation is calculated with Monte Carlo methods, which can be compared with the conservative and progressive bound presented here. Some thresholds regarding to the research questions are also demonstrated.

A. Parameters

The typical parameters of a household load profile are concluded from smart metering data in a Dutch distribution system operator [5]. The key parameters for the case study are listed below:

- \( \mu_m = 0.464\text{ kW} \), \( \mu_{\text{max}} = 1.19\text{ kW} \)
- \( \sigma_{\text{max}} = 1\text{ kW} \)
- \( T = 1\text{ hr} \)
- \( n_i = N_C = 100...500, i \in H \)
- \( |S| = 1...10 \)
- \( \lambda = 0.5039, \phi_d = 1.2075, \lambda_d = 1.4061 \)
- \( \tau_0 = 0.02 \)

Along with \( \lambda, \phi_d \) and \( \lambda_d \), the cumulative distribution function of the user preference factor is shown in Fig. 4. According to smart metering data, it can be assumed that each \( X_{i,j}(t) \) approximately follows log-normal distribution [10].

To simplify the problem without losing generality, the case study assumes that each \( n_i \) for \( i \in H \) has the same value \( N_C \) (every household PoC has the same number of customers). And the case study considers \( N_C \) varying from 100 to 500. As the number of measured PoCs, \( |S| \) (cardinality of \( S \)) varies from one to ten.

The absolute estimation error range varies significantly for different \( N_C \). To make the error ranges comparable, in the case study the relative error range \( \alpha'_i \) is defined with respect to the maximum active power, shown in (38).

\[ \alpha'_i = \frac{|\alpha_i|_{\text{max}}}{n_i \mu_{\text{max}}} \]  
(38)

B. Validity of Central Limit Theorem

To evaluate the error bound of load estimation, this paper assumes that the real-time load at each PoC follows normal distribution, for which theoretically the \( \pm 3\sigma \) bound has a confidence level of 0.997. However, central limit theorem is only valid when the number of customers per PoC is sufficiently large. Therefore, a Monte Carlo simulation is performed to determine the least number of customers per PoC which could ensure the validity of central limit theorem.

Fig. 5 shows the confidence level for \( \pm 3\sigma \) bound, under the actual distribution with different numbers of customers. The confidence level exceeds 0.995, when the number of customers is larger than 50. Therefore, this paper suggests that 50 is the
least number of customers per PoC which could ensure the validity of the error bound approach.

C. Conservative versus Progressive Bound

Fig. 6 shows the real error range from Monte Carlo simulation, and the calculated conservative and progressive bound for different number of customers, with one measured PoC.

For Method 1, the conservative bound is always higher than the real error range with large margin, while the progressive bound is very close to the real error range with almost no risk of being lower than that. Therefore, the paper suggests to use progressive bound for error evaluation.

For Method 2, the conservative bound is also higher than the real error range but with relatively smaller margin. Meanwhile, the progressive bound can be lower than the real error range when the number of customers is smaller than 250. In this case, the choice between conservative and progressive bound depends on whether robustness or performance is preferred.

D. Method 1 versus Method 2

Fig. 7 shows an overview of the real error range for both estimation methods with all the possible combination of \( N_C \) and \(|S|\). It is obvious that the error of Method 2 is always smaller than Method 1, regardless of \( N_C \) or \(|S|\). Thus, when both methods are available to the grid operator, Method 2 (estimation based on annual bill) is always preferred. The reason for this is that estimation based on annual bill internally considers the number of customers, and it also takes individual user preference into account.

E. Number of Measured PoCs

Fig. 8 shows the relation between the real error range and the number of measured PoCs when \( N_C = 300 \). Regarding to research question 3, the error range decreases significantly with measuring more PoCs, when the number of measured PoCs is low. However, with more PoCs measured, installing measurement devices further has less effect.

If the error decrease is less than 1% when further increasing the number of measured PoCs, the paper considers that the error limit is reached. Fig. 9 shows the number of measured PoCs when the error limit is reached for both methods. Usually the error limit is reached when measuring three PoCs, while sometimes this number for Method 1 is larger than that for Method 2.

When the error limit is reached, although the load estimation error cannot be decreased significantly by further increasing measurement points, it needs to be emphasized that increasing measurement points can still improve the performance of the state estimator, because more real measurement data are available.

VI. CONCLUSIONS

To consider the issue that the measurements in MV grids are less than redundant for the state estimator, load estimation has been proposed as pseudo measurements [1]–[3]. However, load estimation should provide not only the values of real-time
Fig. 8. Real error range for different number of measured PoCs, $N_C = 300$

Fig. 9. Number of measured PoCs when the error limit is reached

loads, but also their error range, which is very important for the state estimator. The paper presents a statistical analysis on the error ranges of two commonly proposed load estimation methods for household loads.

Before analyzing the errors, the paper builds a household load profile considering individual user preference. Two load estimation methods, estimation based on number of customers (Method 1) and estimation based on annual bill (Method 2), are introduced.

The estimation errors are divided into system error (representing proportion factors) and random error (representing the diversity of stochastic loads and measurement inaccuracies). For both estimation methods, conservative and progressive bounds are presented to evaluate the error range, using central limit theorem.

A case study with field data from a Dutch distribution system operator demonstrates the feasibility of error evaluation and the comparison of different errors. Generally, the progressive bound should be applied when Method 1 is used. For Method 2, the choice between conservative and progressive bound depends on whether robustness or performance is preferred. The case study also illustrates that Method 2 is more adequate than Method 1 in any situation, and normally the error limit is reached when the number of measured PoCs is three or higher. It should be noted that these results are only valid for the case study. However, the approach described in the paper is analytically precise and applicable.

REFERENCES