



**HAL**  
open science

# Advanced load modelling techniques for state estimation on distribution networks with multiple distributed generators

Weicong Kong, David W. Wang, Colin Foote, Graham Ault, Andrea Michiorri, Robert Currie

## ► To cite this version:

Weicong Kong, David W. Wang, Colin Foote, Graham Ault, Andrea Michiorri, et al.. Advanced load modelling techniques for state estimation on distribution networks with multiple distributed generators. 17th Power Systems Computation Conference, Aug 2011, Stockholm, Sweden. hal-01441892

**HAL Id: hal-01441892**

**<https://hal.science/hal-01441892>**

Submitted on 20 Jan 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ADVANCED LOAD MODELLING TECHNIQUES FOR STATE ESTIMATION ON DISTRIBUTION NETWORKS WITH MULTIPLE DISTRIBUTED GENERATORS

Weicong Kong  
University of Strathclyde  
weicong.kong@eee.strath.ac.uk

David Wang  
Smarter Grid Solutions  
david.wang@smartergridsolutions.com

Colin Foote  
Smarter Grid Solutions  
colin.foote@smartergridsolutions.com

Graham Ault  
University of Strathclyde  
g.ault@eee.strath.ac.uk

Andrea Michiorri  
Smarter Grid Solutions  
andrea.michiorri@smartergridsolutions.com

Robert Currie  
Smarter Grid Solutions  
robert.currie@smartergridsolutions.com

**Abstract** – The paper compares a variety of modelling methods of different complexity to improve the accuracy of pseudo-measurements used for state estimation on distribution networks. The pseudo-measurements are required due to the lack of real-time measurements. However, pseudo-measurements of low accuracy increase the error of the resultant state estimates, degrading the confidence in their use for control applications. The solution is to adopt advanced modelling methods to produce pseudo-measurements with higher accuracy. The candidate methods adopted here are the method of assumed variance, normal distribution fitting, the correlation method and Gaussian mixture modelling. These methods are tested on a real distribution network incorporating distributed generation. The results show that using more advanced methods normally improves the accuracy of state estimates but on some occasions the improvement is not significant and sometimes the accuracy can become worse. Moreover, false large percentage errors in power flow estimates caused by distributed generation have been observed, and this would give wrong indication in ideal locations to add real-time measurements.

**Keywords:** *State Estimation, Distribution Network, Pseudo Measurements, Normal Distribution, Gaussian Mixture Model, Correlation.*

## 1 INTRODUCTION

The connection of renewable and other small-scale generators has complicated the operation of distribution networks (DNs), causing bidirectional power flow and affecting voltage levels. As distribution network operators are obliged to maintain a secure network, it is imperative to understand the impact of new generators at all times. The challenges faced in operating DNs are likely to increase further with new low carbon loads like electric vehicles. Weighted least squares state estimation (SE), which has already been applied widely to transmission networks, is an economically effective approach to improve real-time monitoring of DNs with good reliability and accuracy.

However, low measurement redundancy on DNs often hampers the effectiveness of SE from estimating all the state variables (i.e. bus voltage magnitude and angles) on the DN. To improve the observability of DN, additional measurements are required either by installing more measuring devices, or estimating some

power system quantities based on historical data [1]. The measurements added using the latter approach are called pseudo-measurements and as they are derived from historical data or other indicators they have lower accuracy than real-time measurements taken by measuring devices. Due to budget concerns limiting the number of measuring devices on DN, pseudo-measurements are necessary for distribution SE with the consequence of decreasing SE accuracy.

In earlier distribution state estimation work, the errors of pseudo-measurements were simply assumed by the authors [2, 3] to be much greater than those of real-time measurements, while later more advanced methods were introduced to generate more accurate pseudo-measurements. Ghosh *et al* [4] showed that the pseudo-measurements of loads can be modelled more accurately using a beta distribution rather than a normal distribution; however, the beta distribution cannot be used for the weighted least squares based SE. Mantisas *et al* [5] proposed two methods to model pseudo-measurements; one was based on analysing the correlation between real-time measurements taken at substations and load pseudo-measurements, while in another approach the accuracy of the load pseudo-measurements was improved through Gaussian Mixture Models (GMM). The results showed that both approaches improved the SE accuracy on a 14-bus distribution network to a similar degree. The GMM was further investigated in [6] to produce pseudo-measurements of 55 loads within a generic 95-bus distribution network. In [7], an artificial neural network (ANN) algorithm was adopted to generate pseudo-measurements that temporarily replace the real-time measurements lost during contingent events. However, the pseudo-measurements obtained this way require the prior knowledge of the real-time measurements before the fault rather than a historical load profile. Despite advanced methods being used, there was no mention about how much improvement in SE accuracy can be achieved compared with using simpler methods.

In this paper four pseudo-measurement modelling methods are compared. These are a method based on simple assumptions, normal distribution fitting, more complicated correlation method and GMM. As a result the incentives for adopting more advanced methods are

shown. The four methods are applied to a real distribution network with high renewable generation penetration. The paper is structured as follows: Section 2 explains the fundamental theory of SE while Section 3 describes the four pseudo-measurement modelling methods. The results of SE on the distribution network based on the modelling methods are shown in Section 4. The conclusions are drawn in Section 5.

## 2 STATE ESTIMATION

Each measurement input in SE would have certain degree of errors that can be expressed using the following equation:

(1)

where  $\mathbf{e}$  is a measurement error vector,  $\mathbf{z}$  is a measurement vector,  $\mathbf{x}$  is a system state variables vector, and  $\mathbf{h}$  is a nonlinear measurement function vector.

In the WLS method [8],  $\mathbf{x}$  is calculated that minimises the square of the error vector  $\mathbf{e}$ , the objective function is formulated as follows:

(2)

where  $\mathbf{R}$  is a diagonal error covariance matrix that gives the weighting to each measurement error based on the variance  $\sigma^2$  of the measurements:

(3)

where  $m$  is the total number of measurements used for SE. The state variable vector  $\mathbf{x}$  can be obtained by solving the following linearised derivative of (2) iteratively:

(4)

(5)

(6)

where  $k$  is an iteration number index. In this paper, the WLS method with an augmented matrix is adopted [9]. The main modification is to add two constraints to the objective function (2):

(7)

where  $\mathbf{c}(\mathbf{x})$  is the measurement function vector for the zero power injection measurements on the buses which have neither generators nor loads connected to them.  $\mathbf{r}$  is a residue vector. Equation (4) is 'augmented' while considering the two constraints:

(8)

(9)

where  $\boldsymbol{\lambda}$  and  $\boldsymbol{\gamma}$  are the Lagrange multiplier vectors. The matrices in (8) are in better conditions (i.e. less singular) than those in (2). As a result solving (8) reduces the round-off errors introduced in the iterative process and the risk of solution divergence [9, 10].

As shown in (1) to (9), two inputs are required for each measurement in the WLS based methods: (a) the measurement value and (b) its variance used for computing the weighting matrix  $\mathbf{R}^{-1}$ . Consequently, the pseudo-measurement modelling methods aim at improving the accuracy of these two values.

## 3 PSEUDO-MEASUREMENTS MODELLING METHODS

Given historical load data, the probability density function (PDF) of a load for a specific time period can be computed as shown as the histograms in Figure 1. The methods of assumed variance, normal distribution fitting and Gaussian mixture modelling utilise the PDF to model the pseudo-measurement of the load. In each case the PDF is characterised by one or more values of a mean  $\mu$  and a variance  $\sigma^2$ .

### 3.1 Method of Assumed Variance

Given a historical load profile, the method of assumed variance (AV) finds the measurement value  $\mu$ :

(10)

where  $\alpha_i$ , is the  $i^{th}$  sample of the random variable demand  $\boldsymbol{\alpha}$  and  $n$  is the size of samples taken from the historical data. In AV the standard deviation  $\sigma$  of the pseudo-measurement is simply calculated based on an assumed arbitrary percentage  $\delta$ , normally between 20% and 50%:

(11)

The assumed percentage is fixed for all the load pseudo-measurements regardless of the time and location of the load. In this paper  $\delta$  is equal to 50%.

### 3.2 Normal Distribution Fitting

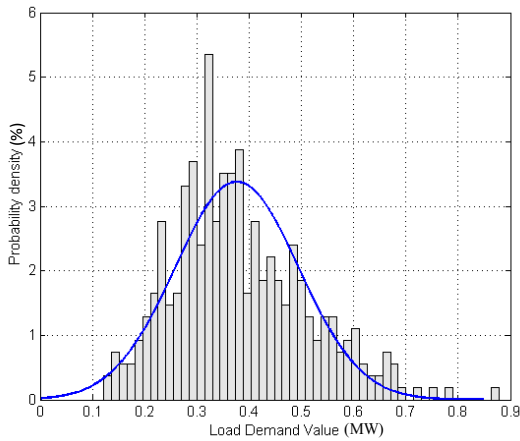
Similar to the AV method, the same measurement value  $\mu$  is obtained using (10). However, the variance is calculated using the conventional statistics equation:

(12)

Figure 1 shows the normal distribution model obtained using the NDF method. It is obvious that in

this example the normal distribution model does not closely mimic the original PDF.

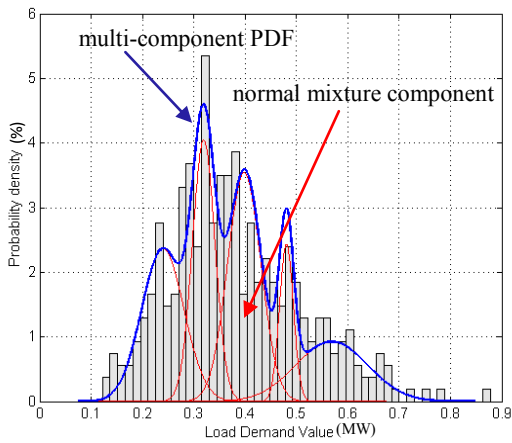
(15)



**Figure 1:** Original load probability density function and the normal distribution model

### 3.3 Gaussian Mixture Modelling

Unlike NDF, Gaussian mixture modelling calculates a multi-component PDF that consists of multiple normally distributed sub-PDFs (mixture components), as depicted in Figure 2.



**Figure 2:** Gaussian mixture model

The multi-component PDF  $f(x)$  is expressed as:

(13)

where  $K$  is the number of mixture components,  $w_k$  is the weight of  $k$  mixture component.  $\mu_k$  is the mean of  $k$  mixture component given by:

$$\mu_k = \frac{\sum_{i=1}^n x_i \cdot f(x_i)}{\sum_{i=1}^n f(x_i)} \quad (14)$$

where  $\mu_k$  is the mean of  $k$  normal mixture component, and

The pseudo-measurement value  $\mu$  is equal to the mean of the normal mixture component closest to the estimated demand, while  $\sigma^2$  is set to the variance of the normal mixture component.

### 3.4 Correlation Method

The method of correlation seeks the linear numerical relationship between a power system quantity  $X$  that is real-time measured and a load  $Y$  by analysing the historical data:

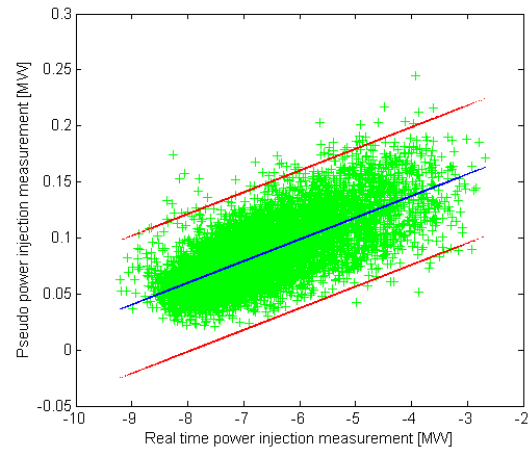
(16)

where  $\mu_x$  and  $\mu_y$  are the means of  $X$  and  $Y$ . In the error covariance matrix  $R$ , the element that corresponds to the real measurement of  $X$  and the pseudo-measurement of  $Y$  is set equal to the covariance calculated.

The covariance coefficient, which has a range between -1 and 1, is then derived:

(17)

where  $\sigma_x$  and  $\sigma_y$  are the standard deviation of  $X$  and  $Y$  respectively. The closer the correlation coefficient to its extreme values, the more linear is the relationship between  $X$  and  $Y$ .



**Figure 3:** Correlation between a real-time measurement and pseudo-measurement

Regression analysis is applied considering the covariance coefficient to find out the degree of dependency between the non-monitored load and the real measurement. As depicted in Figure 3, the regression line in the middle is drawn indicating the linear relationship between the real measurement  $X$  and load  $Y$ . The value of the load pseudo-measurement of therefore lies on the line and depends on the value of the real-time measurement. The interval between the regression line and the one of the boundaries indicates the standard deviation  $\sigma$  of the pseudo-measurement.

## 4 CASE STUDY AND RESULTS

### 4.1 Distribution Network of Orkney

The 33 kV distribution network on the Orkney Islands in the UK is depicted in Figure 4. There are 16 loads connected to the 33 kV network via 33/11 kV transformers. Due to abundant wind resources, the DN has one of the highest levels of wind generation penetration in the UK. Most of the renewable generators are connected on the same branches as loads. As the total generation exceeds the total demand on Orkney, the power is exported from the DN to the mainland transmission network via two sub-sea cables (lines 68 and 78). It is assumed here that all 16 loads are unmonitored and modelled as pseudo-measurements.

### 4.2 Historical Load Data and Result Validation

Generic half-hourly historical load profiles for 2008 and 2009 [11] were used to conduct the study. The load data in 2008 was used for modelling the pseudo-measurements. The load data in 2009 was regarded as the true demand values and used for running load flow analysis to compare the results with those of SE using the pseudo-measurements at the same time stamp. In this study an hourly interval between each time stamp is considered and the period of study is confined in the winter season between December to January.

The estimated error of state variables at time step  $t$  is calculated:

$$(18)$$

Where  $s_{SE}$  and  $s_{PF}$  are the state variable values obtained from the SE and power flow methods respectively. The average estimated error over a period can also be calculated by dividing the aggregated estimated errors by the number of time stamps in the considered period.

### 4.3 Results

The average error of using the four methods to estimate (a) voltage magnitude and (b) voltage angle regardless the bus locations is shown in Figure 5. Overall, the method of correlation produced the most accurate pseudo-measurements, followed by the GMM. However, the improvement in the voltage magnitude accuracy is not great compared with the AV and NDF methods, as all the methods can yield average accurate voltage magnitude estimates with less than 0.05% error. The accuracy of the voltage angle estimates is worse than that of voltage magnitude estimates, while more remarkable improvement has been shown using the correlation method and GMM than the AV method, giving clear incentive to adopt more advanced modeling methods.

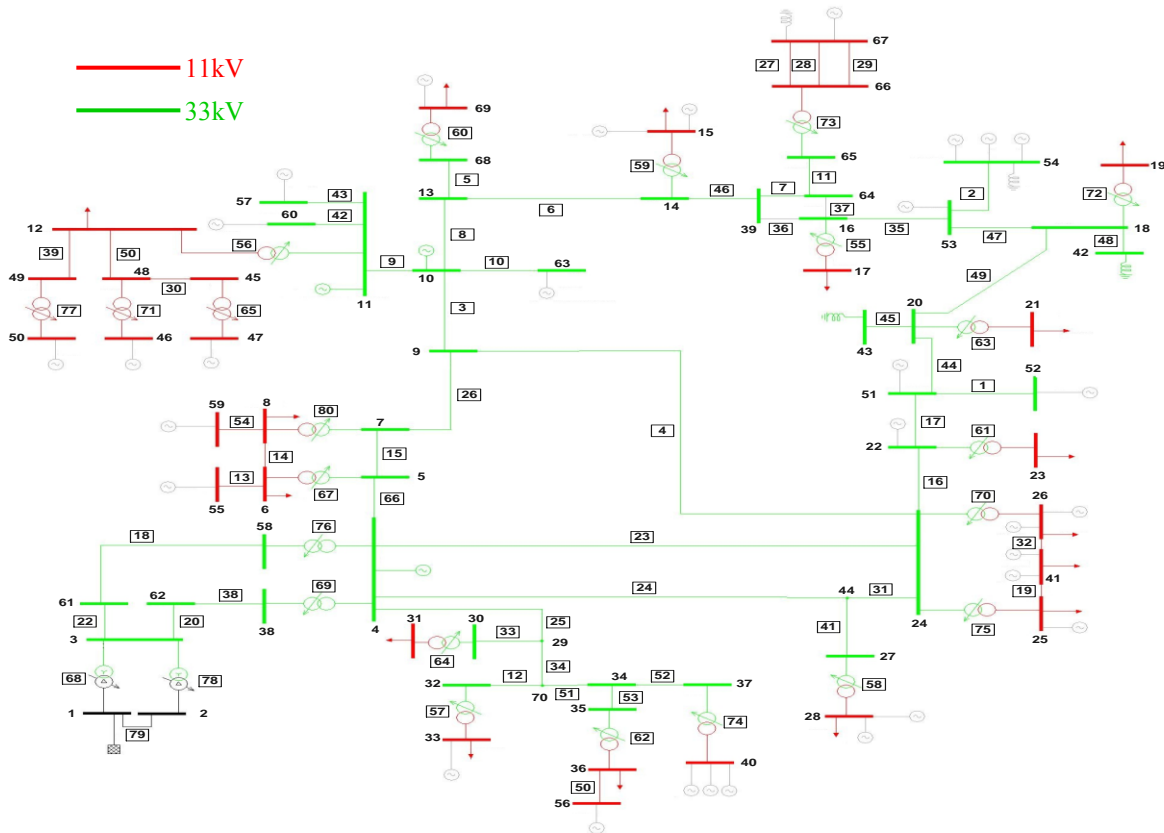
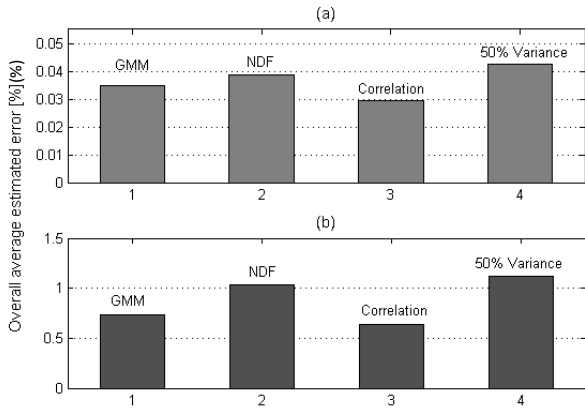


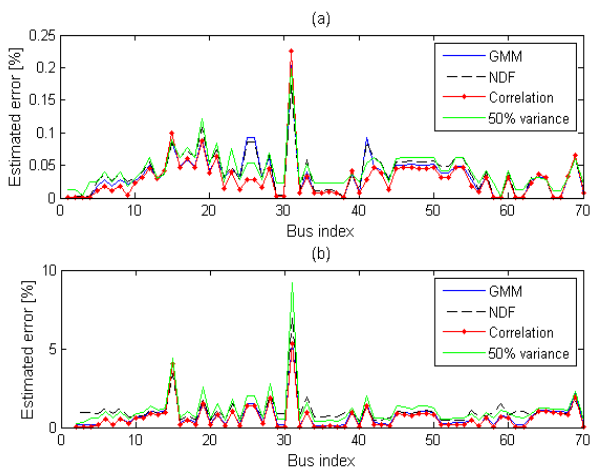
Figure 4: 70-bus distribution network on Orkney Islands

It can be observed from Figure 2 that GMM matches a historical load behavior closer than the NF and AV, therefore giving higher accuracy of the pseudo-measurements and resultant state estimates. As AV, NF and GMM do not have any dependency with real-time measurements, the accuracy of these models is prone to uncertainties, such as weather and temperature, that could render significant difference between the historical and real-time loads. Correlation method, on the other hand, investigates the historical relationship between electrical quantities and real-time measurements and introduces the flexibility of the model to reflect the real-time network conditions.



**Figure 5:** Average percentage error of (a) voltage magnitude (b) voltage angle state estimates regardless bus locations

Figure 6 shows the average error of (a) voltage magnitudes and (b) voltage angles at different buses. The estimated error is highly dependent on the bus locations, and despite the ranking shown in Figure 5, the advanced methods may not always be preferable, such as bus 31 where the correlation method produced the most inaccurate state estimate in average, and for bus 26 and 27 the GMM becomes the worst method.

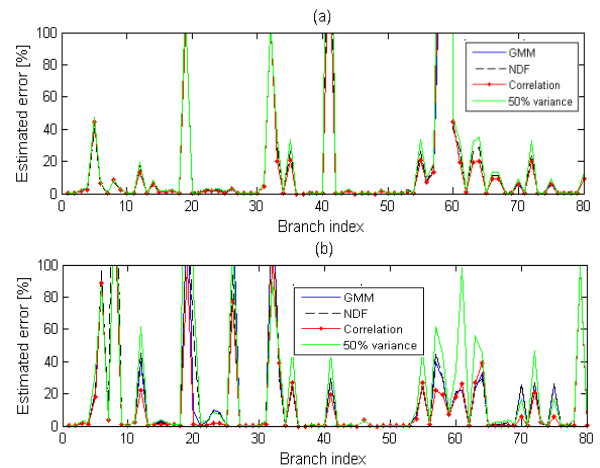


**Figure 6:** Estimated average error percentage of (a) voltage magnitude (b) voltage angle state variables using different modelling methods

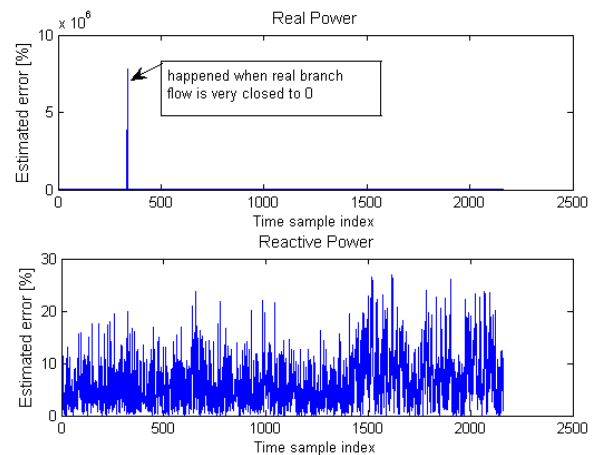
Figure 7 depicts the average estimated errors of branch (a) real and (b) reactive power flow estimates. As expected in overall using the method of correlation

yields the most accurate result. However, some error spikes that are over 100% were observed and none of the four methods seemed to be able to prevent these spikes. Further analysis showed that these error spikes happened on the branches that have both load and generation connected downstream, such as branch 59.

Figure 8 shows the estimated real and reactive power errors on branch 59 for the total 2160 hours using the method of correlation. At hour 336 the real generation output at bus 15 is very close to the demand at the same bus, resulting in almost zero real branch flows in branch 59. Using (8), even slight error in the pseudo-measurement would give an extremely large error percentage that significantly increases the average estimated error in Figure 7. Therefore using the accuracy index in percentage could give false indication in the ideal locations to add real-time measurements to reduce the errors, such as the practice conducted in [12]. As there will be more renewable generation connection at various locations on Orkney, more branches will occasionally experience nearly zero power flows in future. Nearly zero reactive power flow in branch 59 did not happen in the period of analysis, as shown in Figure 8, and the error was between 0 and 25%.



**Figure 7:** Estimated average error percentage of (a) real power flow (b) reactive power flow estimates using different modelling methods



**Figure 8:** Estimated error percentage of real and reactive power flow estimates on branch 59 recorded in 2160 hours.

State estimation is essential for many active network control schemes such as real-time constraint management. It is important for these schemes to issue accurate control instructions by considering the potential degree of errors in their inputs generated by state estimation. Therefore, the percentage errors may not only lead to incorrect decisions on locations for installing new measuring devices, but also have adverse impact on the control accuracy and reliability.

One solution is to measure the errors in absolute values instead of percentages. As shown in Figure 9, the absolute errors of the real power flow estimates in branch 59 for the 2160 hours were all within 0.06 MW, eliminating the spike observed in Figure 8.

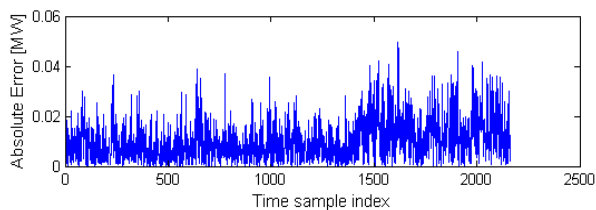


Figure 9: Estimated absolute error of real power flow estimates on branch 59 recorded in 2160 hours.

## 5 CONCLUSION

In this paper, four modelling methods were applied to improve the accuracy of pseudo-measurements used for the weighted least squares based state estimation method. The modelling method of assumed variance and the normal distribution fitting approach utilised conventional statistics to model the pseudo-measurements as normally distributed probability density functions. In the more complicated Gaussian mixture method, the pseudo-measurements were modelled as a multi-component function that consists of multiple sub-models that are normally distributed. The last method was based on analysing the correlation between real-time measurements and pseudo-measurements.

A real distribution network was used for the study. The results showed that in overall the errors of the state estimates were the smallest when the pseudo-measurements were modelled using the method of correlation, followed by Gaussian mixture modelling. Although the method of assumed variance and the normal distribution fitting were in overall the worst methods in this study, for some state estimates the improvement in accuracy was not significant using correlation or Gaussian modelling compared to the two simpler methods. Even in few cases using the method of assumed variance and the normal distribution fitting yielded more accurate result than the correlation method and Gaussian mixture models.

Extremely high error percentages have been observed on some power flow estimates of the branches that connect both generators and loads. When the branches experienced nearly zero power flow due to generation output closely matching demand, even slight error in the pseudo-measurement resulted in large percentage error

of the branch flow estimates. In this condition the accuracy indicator expressed as a percentage would become inappropriate and give false information about the potential locations to install real-time measurements to reduce state estimation errors.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the support and assistance of Scottish & Southern Energy, particularly David MacLeman and Martin Lee, in providing network information and advising on the practicalities of implementation.

## REFERENCES

- [1] M. E. Baran, "Challenges in State Estimation on Distribution Systems", IEEE Power Engineering Society Summer Meeting, pp 1-5, 2001.
- [2] K. Li, "State Estimation for Power Distribution System and Measurement Impacts", IEEE Trans. on Power Systems, Vol. 11, No. 2, pp. 911-916, May 1996.
- [3] C. N. Lu, J. H. Teng and W.-H. E. Liu, "Distribution State Estimation". IEEE Trans. on Power Systems, Vol. 10, No. 1, pp 229-240, 1995.
- [4] A. K. Ghosh, D. L. Lubkeman and R. H. Jones, "Load Modeling for Distribution Circuit State Estimation", IEEE Trans. on Power Delivery, Vol. 12, No. 2, pp 999-1005, 1997.
- [5] E. Mantis et al, "Modelling of Pseudo-measurements for Distribution System State Estimation", CIRED, pp 1-4, 2008.
- [6] R. Singh, B. C. Pal, and R. A. Jabr, "Distribution System State Estimation through Gaussian Mixture Model of the Load as Pseudo-measurement", IET Generation, Transmission and Distribution, Vol. 4, No. 1, pp 50-59, 2010.
- [7] M. B. Do Coutto Filho, J. C. S. de Souza and M. T. Schilling, "Generating High Quality Pseudo-Measurements to Keep State Estimation Capability", Power Tech. 2007.
- [8] A. Abur and A. G. Exposito, "Power System State Estimation: Theory and Implementation", Marcel Dekker Inc., 2004.
- [9] A. Gjelsvik, S. Aam, and L. Holten, "Hachtel's Augmented Matrix Method - A Rapid Method Improving Numerical Stability in Power System Static State Estimation", IEEE Trans. on Power Apparatus and Systems, Vol. 104, No. 11, pp 2987-2993, 1985.
- [10] L. Holten et al, "Comparison of Different Methods for State Estimation", IEEE Trans. on Power Systems, Vol. 3, No. 4, pp 1798-1806, 1988.
- [11] Electric Reliability Council of Texas, "Hourly Load Data Archives", 2010, [Online], retrieved from: <http://www.ercot.com>, at 15/Aug/2010.
- [12] O. Chilard et al, "Distribution State Estimation Based on Voltage State Variables: Assessment of Results and Limitations", CIRED, 2009.