MODELLING OF PSEUDO-MEASUREMENTS FOR DISTRIBUTION SYSTEM STATE ESTIMATION

Efthymios MANITSAS
Imperial College London – United Kingdom
e.manitsas06@ic.ac.uk

Ravindra SINGH
Imperial College London – United Kingdom
ravindra.singh@ic.ac.uk

Bikash PAL
Imperial College London – United Kingdom
b.pal@ic.ac.uk

Goran STRBAC
Imperial College London – United Kingdom
g.strbac@ic.ac.uk

ABSTRACT

The main particularity of distribution system state estimation is the lack of real-time measurements. In order to establish the state estimation function, pseudo-measurements need to be introduced. As there is generalised uncertainty in the power demand, the load characteristics can be utilised to appropriately model the pseudo-measurements.
This paper proposes two new approaches for modelling pseudo-measurements for the purpose of distribution system state estimation; the first is based on correlation and the second is based on load probability density functions.

INTRODUCTION

In recent years, a lot of changes occurred in the way that every power system operates. Large penetration of small scale generation - widely known by the term distributed generation - is now a common feature of many distribution networks.
Active distribution networks seem to be the panacea to all the problems caused by distributed generation and it is widely accepted that active management of distribution network will increase the capacity of distributed generation that the distribution network can accommodate. A necessary function for the materialisation of active distribution networks is the state estimation function.
One of the main particularities of Distribution System State Estimation (DSSE) is the lack of real-time measurements. In distribution, real-time measurements are typically found at the main substations; lines and loads are not usually monitored - not even low voltage substations. There is a generalised uncertainty about the power demand conditions and the line loading, thus pseudo-measurements need to be introduced in order that the state estimation mathematical models can be established and a unique solution can be obtained.
In order for DSSE to be an effective and useful tool, acceptable accuracy has to be achieved. For this to happen, real-time measurements need to be introduced in addition to pseudo-measurements. However, the fact that the majority of measurements used in state estimation are pseudo-measurements reveals the paramount importance of their appropriate modelling, so that they represent the network conditions as realistic as possible.
It is natural to model pseudo-measurements through normal distribution because of its compatibility to Weighted Least-Squares (WLS) estimation based on the maximum likelihood theory. However, the normal distribution assumptions of load profiles, adopted in many papers, do not reflect a realistic situation. In [1], Seppala has suggested log-normal distribution models which were verified from hourly load measurement data obtained from a Finnish load research project. Ghosh et al. [2] have investigated this issue further through load correlation coefficients using diversity factors. They have validated various models such as normal, log-normal and beta distribution through chi-square goodness of fit test.
In [3], an Artificial Neural Network (ANN) scheme is used for producing pseudo-measurements capable of describing the system operating conditions and in [4], the load profiles of non-monitored consumers are determined by using Linear Programming (LP) and taking advantage of typical load profiles and data from metered consumers. In addition, load allocation techniques based on a fuzzy state estimator and l_1-estimation are demonstrated in [5] and [6], respectively. Furthermore, although not specifically targeting to distribution networks, reference [7] is introducing a Probabilistic Autoassociative Memory (PAM) model for tackling problems with missing measurements. Finally, [8] presents an interesting DSSE approach with non-Gaussian, statistically correlated variables.
This paper proposes two new approaches in modelling pseudo-measurements for DSSE.
The first approach entails the calculation of correlation coefficients between real-time measurements at the main substation and normally non-monitored electrical quantities at other buses and lines and the application of regression analysis.
In the second approach, the variability in load distribution is modelled through a Gaussian Mixture Model (GMM) approximation. The advantage of GMM approach is that different types of load distributions can be fairly represented as a convex combination of several normal distributions with their respective means and variances.
In both cases, historical data or temporary measurements are used and the improvement introduced in the accuracy of DSSE is assessed.

STATE ESTIMATION

The general state estimation problem is mathematically given by:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e} \qquad (1)$$

where $\boldsymbol{z}$ is the measurement vector, $\boldsymbol{x}$ is the state vector, $\boldsymbol{h}$ is the measurement function vector and $\boldsymbol{e}$ is the measurement random errors vector.
For the purposes of this paper, the WLS technique was used. This involves finding the vector $\boldsymbol{x}$ that minimizes the objective function $J(\boldsymbol{x})$, defined as:

$$J(x) = [\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})]^T \boldsymbol{R}^{-1} [\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] \qquad (2)$$

where $\boldsymbol{e} \sim N(0, \boldsymbol{R})$ is zero mean Gaussian noise with error covariance matrix defined as:

$$\boldsymbol{R} = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2) \qquad (3)$$

The derivative of this objective function is linearised and iteratively solved to obtain the update in estimate of the states as follows:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{G}^{-1}(\boldsymbol{x}_k)\boldsymbol{H}^T(\boldsymbol{x}_k)\boldsymbol{R}^{-1}[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}_k)] \qquad (4)$$

where

$$\boldsymbol{G}(\boldsymbol{x}_k) = \boldsymbol{H}^T(\boldsymbol{x}_k)\boldsymbol{R}^{-1}\boldsymbol{H}(\boldsymbol{x}_k) \text{ and } \boldsymbol{H}(\boldsymbol{x}) = \frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}} \qquad (5), (6)$$

are the system Gain matrix and the measurement Jacobian matrix, respectively.

METHODOLOGY

A. Correlation Approach

Let X and Y be two random variables. The joint second moment about the means of μ_X and μ_Y is the covariance of X and Y defined as:

$$cov(X,Y) = \mathrm{E}[(X - \mu_X)(Y - \mu_Y)] \qquad (7)$$

Covariance is a measure of the degree of linear relationship between variables X and Y. The normalised covariance or correlation coefficient is defined as:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \qquad (8)$$

where σ_X and σ_Y are the standard deviations of random variables X and Y, respectively. The correlation coefficient takes values between -1 and +1. The closer the correlation coefficient is to its extreme values, the more linear is the relationship between variables.
For a set of observations $\{(x_i, y_i), i=1,\ldots,n\}$, the correlation coefficient is given from:

$$\rho_{X,Y} = \frac{\sum_{i=1}^{n}(x_i - \mu_X)(y_i - \mu_Y)}{\sqrt{\sum_{i=1}^{n}(x_i - \mu_X)^2 \sum_{i=1}^{n}(y_i - \mu_Y)^2}} \qquad (9)$$

In the context of DSSE and assuming that measurements are not independent, the non-diagonal elements of the measurement covariance matrix are equal to:

$$R_{ij} = \sigma_i \rho_{ij} \sigma_j \qquad (10)$$

where σ_i is the standard deviations of measurement i, σ_j is the standard deviations of measurement j and ρ_{ij} is their correlation coefficient. Taking advantage of the fact that real-time measurements are usually available only at the source substation and using historical data or temporary measurement devices, correlation coefficients between real-time measurements at the source substation and electrical quantities at buses or lines that are not normally monitored can be calculated.
With the same data, regression analysis can be applied using real-time measurements at the substation as independent variables and electrical quantities at non-monitored buses as dependent variables. Depending on their correlation coefficient, linear or non-linear regression can be used. Along with the regression line, prediction intervals are identified which, in turn, are translated to pseudo-measurement standard deviations.

B. Load Probability Density Function Approach

The variability in the load probability density function (pdf) was modelled using a Gaussian Mixture (GM) model approximation. A GM pdf is a weighted, finite sum of Gaussian pdfs. It is characterized by the number of mixture components and their corresponding weights, means and variances. Since a pdf must be nonnegative and the integral of a pdf over the sample space of the random quantity it represents must evaluate to unity, the mixture weights must be nonnegative and sum of all the weights must equal to one. For the multivariate case, the GM pdf model can be given by:

$$f(\boldsymbol{z} \mid \gamma) = \sum_{i=1}^{M_c} w_i f(\boldsymbol{z} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \qquad (11)$$

where M_c is the number of mixture components, w_i is the weight of i^{th} mixture component, subject to:

$$w_i > 0 \text{ and } \sum_{i=1}^{M_c} w_i = 1 \qquad (12)$$

and γ is chosen from the following set of parameters:

$$\Gamma = \left\{\gamma : \gamma = \{w_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{M_c}\right\} \qquad (13)$$

each member of which defines a GM.
Given a d-dimensional random variable z, mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_i$, the pdf of each mixture component is a normal distribution given by:

$$f(z \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{(2\pi)^{d/2} det(\boldsymbol{\Sigma}_i)^{1/2}} e^{-\frac{1}{2}(z-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(z-\boldsymbol{\mu}_i)} \quad (14)$$

The parameters of the mixture components were obtained using the Expectation Maximisation (EM) algorithm in [9].

CASE STUDY

The performance of the state estimator was evaluated using both pseudo-measurements modelling approaches on a part of the 33kV EDF Energy network shown in Figure 1. The network comprises two radial sections, connected via a normally open circuit breaker. The results corresponding to section 1 (Buses 1, 2 and 3) are presented in this paper.

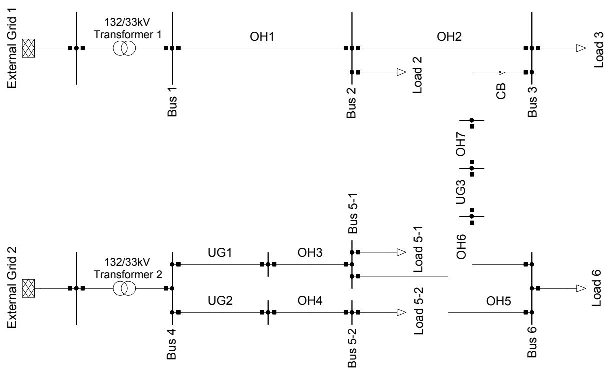

Figure 1: 33kV EDF Energy Network

Twenty four hour, one minute measurement data were obtained from the EDF Energy control centre for a typical winter day. The voltage and current magnitudes at the main substation were used as real-time measurements in state estimation. Voltage angle measurements were created using load flow simulations for comparison purposes. Active and reactive load data at Buses 2 and 3 were used to create pseudo-measurements using the two approaches.
The accuracy of the real-time measurements was assumed to be 1%. The standard deviation of error in the i^{th} real-time measurement was calculated using the following formula:

$$\sigma_i = z_i \frac{accuracy_i}{300} \quad (15)$$

where z_i is the measurement value and σ_i is the standard deviation of the error. The standard deviation of error in pseudo-measurements was derived from the corresponding modelling approach.
In both approaches, 99.7% confidence intervals were used for real-time measurements and pseudo-measurements, which correspond to $\pm 3\sigma$.

Figure 2 demonstrates the concept of the Correlation approach. The red markers represent 1440 (I_{12}, P_2) measurement points, the blue continuous line indicates the linear approximation of I_{12} against P_2, and the blue dotted line shows the 99.7% prediction interval.

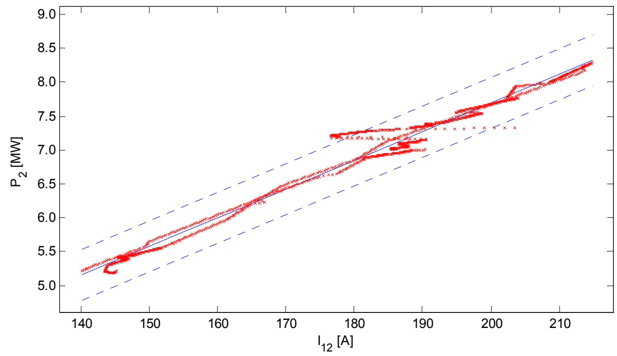

Figure 2: Correlation Approach

Figure 3 demonstrates the concept of the Load Probability Density Function approach. The blue bars represent the original distribution of P_2, the red dotted line shows the individual GMM components, and the black line outlines the GMM pdf. In this case, the load pdf is represented by two Gaussians.

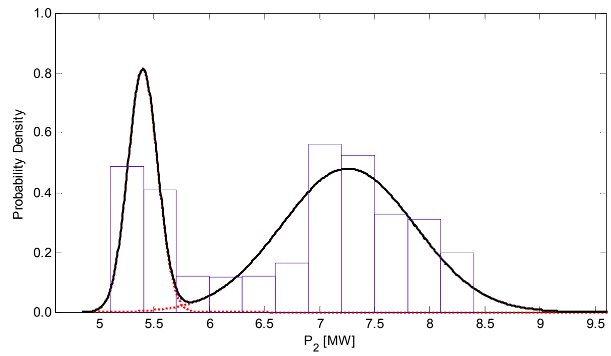

Figure 3: Load Probability Density Function Approach

The voltage magnitude and voltage angle estimates using both approaches are shown in Figures 4 and 5, respectively. The red continuous lines represent the actual measurement, the blue continuous line is the mean value of the estimate and the blue dotted lines are the 99.7% confidence intervals ($\pm 3\sigma$). In all cases, the estimates follow the pattern of the actual measurements.
As real-time measurements were used at the main substation only, the mean value of the voltage magnitude estimate at the remote Bus 3 shows the largest deviation from the actual measurement. However, in all voltage magnitude estimates, the actual measurements lie between the 99.7% estimate confidence bounds, indicating voltage magnitude estimates that can be trusted at all times.

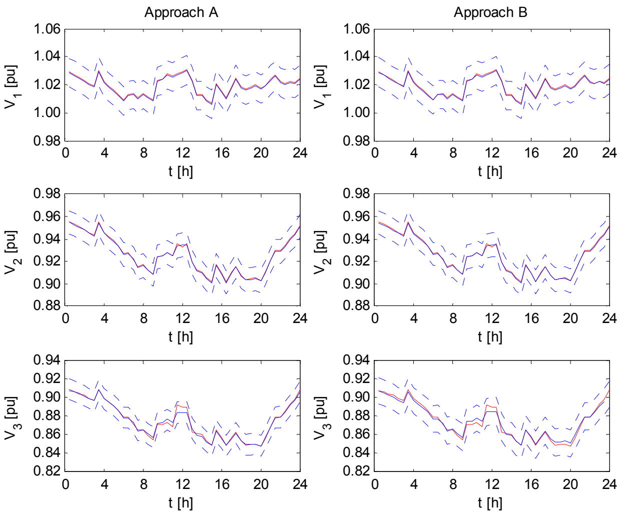

Figure 4: Voltage Magnitude Estimates

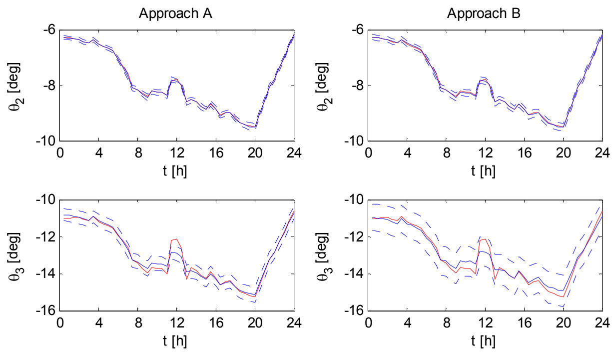

Figure 5: Voltage Angle Estimates

As for the voltage angle estimates, both approaches succeed in providing good estimates at Bus 2. The actual measurements lie between the 99.7% estimate confidence interval at all times using both methodologies. On the other hand, only approach B succeeds in providing good estimates at Bus 3 at all times. Approach A fails to do so around hour 12, where the actual measurements lie outside the 99.7% estimate confidence interval. As shown in Figure 2, this is due to the fact that, as at that time I_{12} is 175-180A, the actual (I_{12}, P_2) measurement points lie outside the 99.7% pseudo-measurement modelling prediction bounds. Increasing the prediction bounds at the pseudo-measurement modelling stage can overcome this shortfall.

CONCLUSION

Two approaches for modelling pseudo-measurements have been presented. The performance of the WLS estimator using the proposed approaches was evaluated on a practical test network. Reliable voltage magnitude and voltage angle estimates were obtained. The performance of the two approaches indicates that they can be benefial for Distribution System State Estimation. As an extension of this work, these techniques can be applied on large distribution networks at lower voltage levels.

Acknowledgment

The authors would like to thank EDF Energy Networks for supporting this project under grant EESC-P05821.

REFERENCES

[1] A. Seppala, 1995, "Statistical Distribution of Customer Load Profiles", *in proc. IEEE Int. Conf. on Energy Management and Power Delivery,* vol. 2, 696-701

[2] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, 1997, "Distribution Circuit State Estimation Using a Probabilistic Approach", *Power Systems, IEEE Transactions on,* vol. 12, no. 1, 45-51.

[3] A. Bernieri, G. Betta, C. Liguiri, and A. Losi, 1996, "Neural Networks and Pseudo-Measurements for Real-Time Monitoring of Distribution Systems", *Instrumentation and Measurement, IEEE Transactions on,* vol. 45, no. 2, 645-650.

[4] D. Gerbec, F. Gubina, and Z. Toroš 2005, "Actual Load Profiles of Consumers without Real Time Metering", *Power Engineering Society General Meeting, 2005. IEEE,* vol. 3, 2578-2582

[5] V. Miranda, V. Pereira, and J.T. Savaira, 2000, "Load Allocation in DMS with a Fuzzy State Estimator", *Power Systems, IEEE Transactions on,* vol. 15, no. 2, 529-534.

[6] J. Wan and K. Miu, 2000, "Load Estimation in Radial Electric Power Distribution Networks Using Limited Measurements", *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on,* vol. 2, 517-520

[7] A.P. Alves Da Silva, V.H. Quintana, and G.K.H. Pang, 1992, "A Probabilistic Associative Memory and its Application to Signal Processing in Electrical Power Systems", *Engng Applic. Artif. Intell.,* vol. 5, no. 4, 309-318.

[8] R. Mínguez, A. Conejo, and A.S. Hadi, 2006, "Non Gaussian State Estimation in Power Systems", *International Conference on Mathematical and Statistical Modeling in Honor of Enrique Castillo.*

[9] A. P. Dempster, N. M. Laird, and D. B. Rubin, 1977, "Maximum Likelihood from Incomplete Data via the EM Algorithm", *J. Royal Statist. Soc. Ser. B,* vol. 39, 1-38.