

Loss-of-Main Monitoring and Detection for Distributed Generations Using Dynamic Principal Component Analysis

Yuanjun Guo, Kang Li, D. M. Laverty

School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, UK
Email: yguo01@qub.ac.uk, k.li@qub.ac.uk, david.laverty@qub.ac.uk

Received December 2013

Abstract

In this paper, our previous work on Principal Component Analysis (PCA) based fault detection method is extended to the dynamic monitoring and detection of loss-of-main in power systems using wide-area synchrophasor measurements. In the previous work, a static PCA model was built and verified to be capable of detecting and extracting system faulty events; however the false alarm rate is high. To address this problem, this paper uses a well-known 'time lag shift' method to include dynamic behavior of the PCA model based on the synchronized measurements from Phasor Measurement Units (PMU), which is named as the Dynamic Principal Component Analysis (DPCA). Compared with the static PCA approach as well as the traditional passive mechanisms of loss-of-main detection, the proposed DPCA procedure describes how the synchrophasors are linearly auto- and cross-correlated, based on conducting the singular value decomposition on the augmented time lagged synchrophasor matrix. Similar to the static PCA method, two statistics, namely T^2 and Q with confidence limits are calculated to form intuitive charts for engineers or operators to monitor the loss-of-main situation in real time. The effectiveness of the proposed methodology is evaluated on the loss-of-main monitoring of a real system, where the historic data are recorded from PMUs installed in several locations in the UK/Ireland power system.

Keywords

Loss-of-Main Detection; Phasor Measurement Units; Big Data; Dynamic Principal Component Analysis

1. Introduction

During the last decade, detecting and monitoring power system events have emerged as a focused and promising research topic, which also forms a key part of the Wide-area Monitoring and Control Systems (WAMCS) [1]. Loss-of-main monitoring and detection is one of the most important and hot issues.

There are two major concerns when dealing with the loss-of-main problem. The first one is the "Big Data" issue. By virtue of deploying a wide range of smart grid infrastructures, the utilities are gathering tremendous volumes of data on a much more granular level than ever before. This enables the utilities to potentially design

better response programs, forecast and control the loads, integrate renewable energy and plan the system operation, etc. However, they are facing new analytic challenges in dealing with the big data, and using them to improve decision-making. A recent report on the big data issue in the power system is presented in [2], in which the big data is defined as increasing large volume, high velocity, increasing variety, or all the three. Moreover, the data volumes in the power industry are growing faster than many power utilities may be prepared to handle [3].

The second concern is the variety and multiformity of the power generations. In UK and Ireland, the power networks operate with greater than 10% of power supplied by distributed generations, especially wind farms. Due to common weather patterns, the wind power will exhibit reasonably strong statistical dependency within a certain regional area. Therefore, a statistically efficient process monitoring and control techniques have been chosen for power system transient analysis, loss-of-main detection and extraction to address the above mentioned concerns. Burke and O'Malley [4] applied principal component analysis (PCA) to the covariance matrix of distributed wind power data from existing Irish wind farms, with eigenvalue analysis generating a lower number of uncorrelated alternative variables. The same idea was also implemented in our previous work [5], but with an extended application in power system islanding detection as well as a fault reconstruction method that have been used to identify the islanding site based on the data collected from synchrophasors in a real system. Similarly, in a recent paper by Barocio *et al.* [6], PCA was used for detecting and extracting unusual or anomalous events from wide-area monitoring data, demonstrating the ability to detect and extract system events based on a transient stability model of the New England Test System.

All these researches adopt the PCA method to reduce the dimensionality of the big data to a certain extent. However, it is worth noting that the PCA works well under three assumptions: 1) data follows a multivariate normal distribution, 2) there exist no autocorrelation among observations, and 3) the variables are stationary (keep constant mean and standard deviation over time). From our previous work in [5], it concluded that static PCA based fault detection method could address the problem of redundant and high dimensional data. However, challenges still exist due to the dynamic, and multi-timescale synchronized measurements, which is the most common situation that undergoes slow and normal process changes in real systems.

In case of processes with small disturbances, the autocorrelation in variables is taken into account, incorporating the time lags of time series during the modelling stage, this extension is called dynamic principal component analysis (DPCA). The purpose of this paper is to show how static PCA tools could be extended for the dynamic multivariate systems and to examine their use for disturbance analysis and loss-of-main detection in the electrical power system. This is achieved by building a dynamic PCA model through singular value decomposition method on the covariance matrix of the time lagged 'trajectory data'. Similar to static PCA, the data is then projected into principal component subspace and residual space, and two statistics are used to detect loss-of-main situations.

This paper is organized as follows. Section 2 introduces the dynamic PCA approach in combination with the power system synchrophasor measurements, including the trajectory data matrix formation as well as the disturbance analysis and extraction by the dynamic PCA technique. Section 3 gives a brief introduction of the real system framework. Section 4 presents the data analysis results according to different variables and PMU combination cases step by step, to give a clear insight of how DPCA works in the application of loss-of-main detection and monitoring. Finally, Section 5 concludes the paper and discusses the future work.

2. Dynamic Principal Component Analysis

2.1. Data Matrix of DPCA

Let matrix \mathbf{X} be a set of historical data made up of n observations from p variables, which is taken from a PMU recordings of a distributed power system dynamic process. When the process is under nominal conditions and operates around an operating point, it can be described by the matrix.

$$X = [X_1 \quad X_2 \quad \cdots \quad X_p]_{(n \times p)} \quad (1)$$

Each column in \mathbf{X} represents an auto-correlated time series, where a current value depends on the past values. In order to include the serial correlation of data, the so called "trajectory matrix" is constructed applying a time lag shift of order w on each of the p columns of the matrix \mathbf{X} . The trajectory matrix lagged from one variable

column is given as follows

$$X_i^w = \begin{bmatrix} X_i(1) & X_i(2) & \cdots & X_i(w) \\ X_i(2) & X_i(3) & \cdots & X_i(w+1) \\ \vdots & \vdots & \ddots & \vdots \\ X_i(n) & X_i(n+1) & \cdots & X_i(n+w-1) \end{bmatrix}_{(n \times w)} \quad (2)$$

therefore, the complete trajectory matrix with p variables is

$$X^w = [X_1^w \quad X_2^w \quad \cdots \quad X_p^w]_{(n \times m)} \quad (3)$$

where $m = pw$. The choice of w is made establishing a compromise between the information content and the statistical confidence. To avoid the situation that some particular variables dominate the modelling process, it is convenient to carry out data standardization on matrix X^w in relation to its means and standard deviations. Thus, the means of X^w are given by

$$\mu_{X^w} = \left[\frac{1}{n} X^{wT} \mathbf{1} \right]^T = [\mu_1 \quad \mu_2 \quad \cdots \quad \mu_m]_{(1 \times m)} \quad (4)$$

where $\mathbf{1} = [1 \quad 1 \quad \cdots \quad 1]^T \in \mathfrak{R}^n$. On the other hand, the covariance matrix of X^w is given by

$$S_{X^w} = \frac{1}{n-1} (X^w - \mathbf{1} \mu_{X^w})^T (X^w - \mathbf{1} \mu_{X^w}) \quad (5)$$

from which the standard deviations can be obtained as follows

$$\hat{X}^w(i, j) = \frac{X^w(i, j) - \mu_{X^w}(j)}{\sigma_{X^w}(j)} \quad (6)$$

where $i = 1, \dots, n$ and $j = 1, \dots, m$. The data formed as trajectory matrix both for off-line model training and for assessing new coming data. Applying PCA to \hat{X}^w is so called DPCA.

2.2. Disturbances Detection via DPCA

The principal component are linear transformation of the original variables, such principal components constitutes a new set of variables Z of dimension $n \times l$ which are uncorrelated to each other. They are obtained by

$$Z = \hat{X}^w V \quad (8)$$

where the orthonormal transformation matrix $V \in \mathfrak{R}^{m \times l}$ is composed of an appropriate selection of l eigenvectors, called loading vectors, associated to the correlation matrix $R = \frac{1}{n-1} \hat{X}^{wT} \hat{X}^w$. The data matrix \hat{X}^w can be expressed as

$$\hat{X}^w = \tilde{\hat{X}}^w + E = ZV^T + E \quad (9)$$

where $\tilde{\hat{X}}^w$ is the information captured by the l principal components and E is the information in the neglected $m-l$ components. So for the detection purpose it is possible to use the Hotelling static T^2 from Z and/or the squared prediction error (SPE) from E . Calculation for T^2 is given by

$$T^2 = ZS^{-1}Z^T \quad (10)$$

where S is the covariance matrix of Z . Finally, a threshold of normal condition from the probability density function of the set of parameters is calculated. In literature [7], for the data set T^2 that follows a beta distribution, the threshold Upper Control Limit (UCL) is given by

$$UCL = \frac{(n-1)^2 \left(\frac{1}{n-l-1} F\left(\frac{\alpha}{2}; l; n-l-1\right) \right)}{n \left(1 + \frac{1}{n-l-1} \right) F\left(\frac{\alpha}{2}; l; n-l-1\right)} \quad (11)$$

where n and l are the dimensions of \mathbf{Z} and α is a level of significance.

A complementary test on the Q statistic, defined as the sum of the squared residual of the selected PCs, is used to measure the variation in the residual space not accounted for by the PCA model.

$$Q = \tilde{\mathbf{X}}^w (I - \mathbf{V}\mathbf{V}^T) \quad (12)$$

which is also interpreted as a measure of the “best fit” of the PCA model, showing the distance between the actual and the predicted data. A suitable threshold for statistical significance of Q is given by [8].

$$Q_i < Q_\alpha = \theta_1 \left[\frac{c_\alpha h_0 \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \quad (13)$$

with

$$\theta_i = \sum_{k=l+1}^m \lambda_k^i, i=1, \dots, m \quad \text{and} \quad h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_1^2}$$

where Q_α is the upper confidence limit for the residual model Q_i with significance level and c_α is the normal deviation corresponding to the upper $(1 - \alpha)$ percentile.

Used together, both the two statistical tests are able to identify whether the event is associated with unusual variability within the principal subspace defined by the selected PCs or by variations not explained by the retained PCs.

3. Synchrophasor Based Loss-of-Main Detection Framework

Li *et al.* [9] developed a novel distributed adaptive real-time learning framework for wide area monitoring of power systems integrated with distributed generations using synchrophasor technology. The framework comprises distributed agents (synchrophasors) for autonomous local condition monitoring and fault detection, and a central unit for generating global view for situation awareness and decision making. The local agent PMUs, which have been deployed by Laverty *et al.* [10] in the system of British/Irish utility networks, were developed and manufactured in conjunction with the Open PMU project in Queen’s University of Belfast (QUB) [11] [12] and Scottish & Southern Energy. An overview of the proposed approach in a physical interpretation is given in **Figure 1**.

As shown in **Figure 1**, several PMUs are installed in different locations of UK/Ireland power grid, including distributed generation level as well as main power stations to collect data including GPS time signatures, voltage magnitude, frequency and phase angle. A reference signal is acquired from a dependable utility site and transmitted to the generator site by means of Internet Protocol (IP) telecommunications [13]. At the generator site, the reference signal is compared to the synchrophasors acquired at the generator terminals. The synchronized measurements with a sampling rate of 10 Hz are then transmitted to QUB server and stored for further analysis.

The time signatures of the synchrophasors are aligned in order to compare and analyse the phasors taken at the same instant. The phase difference between the generator and the reference site are compared against known typical operating conditions. The traditional mechanism of loss-of-main detection is to find the abnormally large phase difference between the generator and the reference site, which is then considered to be in loss-of-main situation. This passive technique faces a difficulty for time-varying operation conditions. Detection based on current single window of power system conditions may fail either to detect the islanding condition or to cause nuisance trip-ping. Therefore, the aforementioned dynamic PCA method along with the incremental learning framework provides a great potential in improving the near real-time monitoring and detection performance.

4. Loss-of-Main Detections Based on Dynamic PCA Cases

4.1. Analysis of Single-Variable and Single-PMU Case

For the case of single-variable and single-PMU, frequency data recorded by the main site PMU was used in order to test the ability of dynamic PCA to detect the abnormal disturbances such as frequency dip caused by an inter-connector tripping. To extract the dynamic relations from the PMU synchrophasors by the use of dynamic PCA, the necessary number of time lags should be determined. A method called Parallel Analysis on the time shifted time could be found in [14].

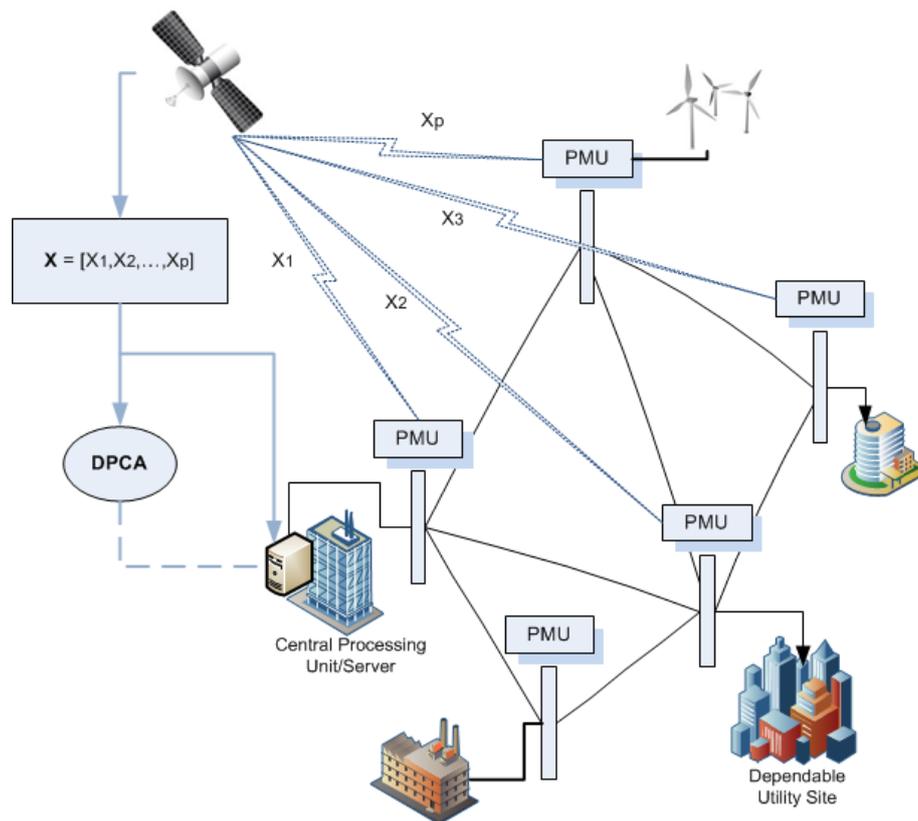


Figure 1. Overview of the proposed synchrophasor based loss-of-main detection scheme using DPCA method.

The trajectory data matrix was set up according to the time lag of 20 seconds with observations of 1000 nominal samples without disturbances to build the dynamic PCA model. A transient event occurred from the sample point of 1306, which was caused by an instantaneous loss of 1 GW power imported to the GB power grid caused by an interconnector tripping between Great Britain (GB) and France, details of the event could be found in [5]. The dynamic PCA model of single-variable and single-PMU data was tested using the data containing the transient event data as well as the nominal data, with the result given in **Figure 2**.

Figure 2 presents the monitoring results using T^2 and Q chart for single-variable and single-PMU data. The red dashed lines in the two charts are the 95% control limits of the two statistics. The enlarged view of T^2 is given in order to clearly show the normal operating region with no false alarms before the transient event happened. Due to the single variable data, only the variance within the principal model can be detected by the T^2 chart, leaving no information in the residual subspace, which explains the clear situation in the Q chart in this case.

4.2. Analysis of Single-Variable and Multi-PMU Case

During the inter-connector failure event, the frequency of the main power system had fallen from 50.08Hz to 49.70 Hz. The result of rate-of-change-of-frequency (ROCOF) exceeded the current ROCOF setting, which falsely triggered an islanding operation on one of the sites and led to further loss-of-main situation. The training and testing data in this case consist of three PMU data, only the frequency variable is used for each PMU. A dynamic PCA model has been built based on the nominal condition data and monitoring result is given in **Figure 3**. It is worth to mention that due to the multi-timescale synchronized measurements corresponding to the different GPS time format and computational precision, the total data number is different when the multi-PMU data are aligned according to the same UTC time instant. Nevertheless, the converted local time is in accordance with that of the single-PMU case, the only difference is the number of sample points.

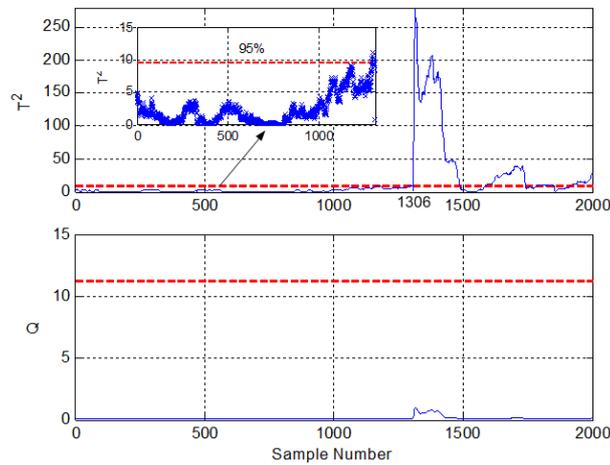


Figure 2. Dynamic PCA based monitoring results using T^2 and Q charts for single-variable and single-PMU data.

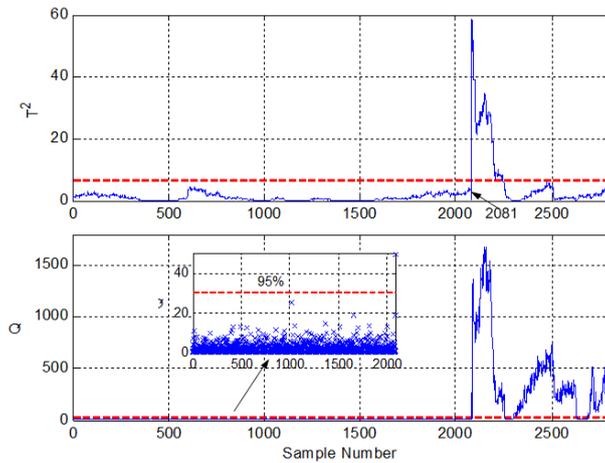


Figure 3. Dynamic PCA based monitoring results using T^2 and Q charts for single-variable and multi-PMU data.

Figure 3 shows the same behaviour in T^2 control chart, which clearly differentiate the frequency dip event with normal operation at the sample point of 2081. The relationship between the different PMU agents can be shown through the Q chart, where two high values exceed the Q control limit. One exceeding value at the same position with T^2 represents the frequency dip event, the other points exceeding the confidence limit continuously reveals a loss-of-main situation. For further analysis of the islanding site identification, fault reconstruction method proposed by Dunia and Qin in [15] was performed in the central unit. The identification result is given in **Figure 4**, which clearly shows the loss-of-main data came from the PMU at the forth site which has the most contribution to the reconstructed fault result.

4.3. Analysis of Multi-Variable and Single-PMU Case

As mentioned above, there are other useful variables included in the synchronized phasor measurements, such as magnitude and phase angles. In order to make the best use of the dynamic PCA approach, the magnitude and frequency were chosen to form the training and test data because of their normal distribution property. **Figure 5** shows the similar behaviour as **Figure 2** except some fluctuations around the 95% confidence limit in T^2 control chart due to the dynamic changes within voltage magnitude. In addition, the small disturbance indicates that the dynamic change did not significantly affect the nominal magnitude. Similar to case A, the single PMU data has no exceeding information shown in the Q chart.

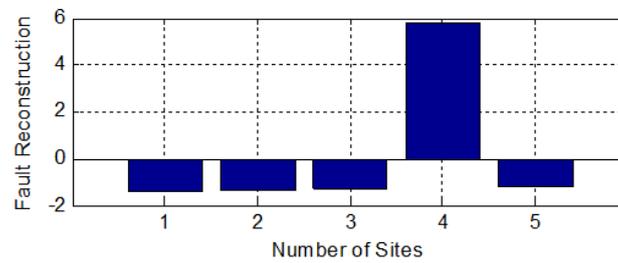


Figure 4. Loss-of-main identification results of 5 different sites.

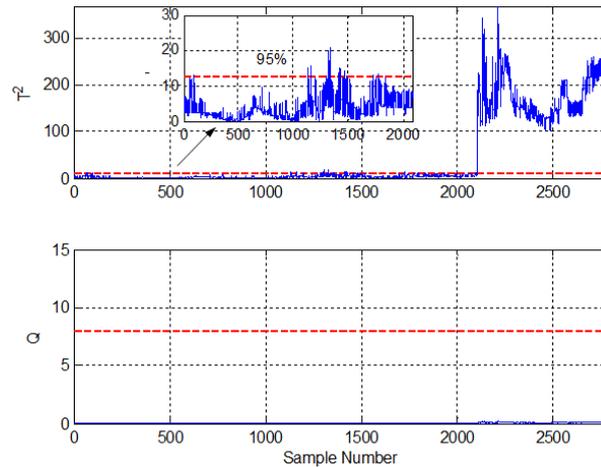


Figure 5. Dynamic PCA based monitoring results using T^2 and Q charts for multi-variable and single-PMU data.

4.4. Analysis of Multi-Variable and Multi-PMU Case

It is important for engineers to monitor the entire grid in order to get the complete information of the status and timely detect any abnormal disturbances. An advantage of applying the PCA-based techniques is to process the high dimensional multi-variate data simultaneously and extract key features of the process. In this section, the trajectory matrix consists of data from 5 different PMUs with magnitude and frequency for each PMU. The dynamic PCA monitoring result is illustrated in **Figure 6**.

This monitoring result integrates the aforementioned different cases, which also shows a good ability to detect both the generation loss trip from T^2 control chart, and the loss-of-main status from Q chart with a neglectable low rate of false alarm. Further analysis by fault reconstruction method can identify the islanding site with a result similar to **Figure 4**.

In summary, the complementary case studies have shown that the proposed dynamic PCA method for the wide-area power system incremental learning framework can offer great potential in fast and dynamical capturing of loss-of-main events while reducing the rate of false alarms and effectively and efficiently deal with the big data issue in the power systems.

5. Conclusions

This paper has shown that the dynamic PCA statistical process control approach is a powerful tool for automatic loss-of-main detection. This method is capable of analysing the auto- and correlations between the variables through decomposition of the covariance matrix of the time-lagged trajectory matrix. Similar as static PCA procedure, the data matrix has been projected into principal subspace and residual space. Two statistics of Hotelling's T^2 and Q are calculated to monitor the variations in the trajectory data. Through the step-by-step analysis of different real system data cases of variables and PMUs, it is clear that most of the abnormal disturbance can be captured by the T^2 control chart, while the loss-of-main event can be intuitively seen through a continuously exceeding trend within the Q control limit. With dynamic PCA approach, the false alarms have been signifi-

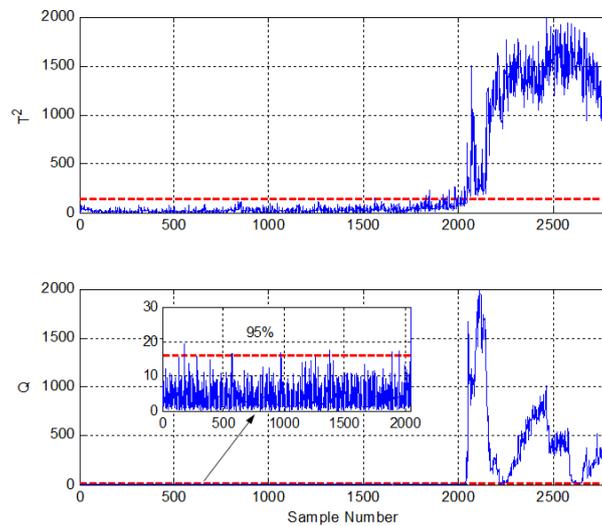


Figure 6. Dynamic PCA based monitoring results using T^2 and Q charts for multi-variable and multi-PMU data.

cantly reduced due to the dynamic characteristic been interpreted well through the properly chosen time lag and number of principal components. The proposed methodology is practically applicable and efficient from the tests on real system cases.

The work however has some limitations such as the dynamic PCA will fail to detect the loss-of-main situation completely when the multi-timescale data are aligned together which may filter out small but important sections of data. More research needs to be done to investigate the relationship between different time lags and the number of principal components under different abnormal power system disturbances. Moreover, dynamic extensions will be applied to more advanced methodologies to make use of the phase angles in the future work to address the non-linearity issues.

Acknowledgements

This work was partially funded by Engineering and Physical Sciences Research Council (EPSRC) under grants EP/L001063/1 and EP/G042594/1. Y Guo would like to thank China Scholarship Council for the financial support.

References

- [1] Kamwa, I., Grondin, R. and Hebert, Y. (2001) Wide-Area Measurement Based Stabilizing Control of Large Power Systems—A Decentralized/Hierarchical Approach. *IEEE Transactions on Power Systems*, **16**, 136-153. <http://dx.doi.org/10.1109/59.910791>
- [2] IBM Softwares (2012) Managing Big Data for Smart Grids and Smart Meters.
- [3] Kezunovic, M., Xie, L. and Grijalva, S. (2013) The Role of Big Data in Improving Power System Operation and Protection. *IEEE of Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP)*, 1-9.
- [4] Burke, D.J. and O'Malley, M.J. (2011) A Study of Principal Component Analysis Applied to Spatially Distributed Wind Power. *IEEE Transactions on Power Systems*, **26**, 2084-2092. <http://dx.doi.org/10.1109/TPWRS.2011.2120632>
- [5] Guo, Y., Li, K. and Laverty, D.M. (2013) A Statistical Process Control Approach for Automatic Anti-Islanding Detection Using Synchrophasors. *IEEE of Power and Energy Society General Meeting*, Vancouver, 21-25 July 2013, 1-5.
- [6] Barocio, E., Pal, B.C., Fabozzi, D. and Thornhill, N.F. (2013) Detection and Visualization of Power System Disturbances Using Principal Component Analysis. *2013 IREP Symposium of Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP)*, 1-10.
- [7] Tracy, N., Young, J. and Mason, R. (1992) Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, **24**.
- [8] Wise, B.M. and Gallagher, N.B. (1996) The Process Chemometrics Approach to Process Monitoring and Fault Detec-

- tion. *Journal of Process Control*, **6**, 329-348. [http://dx.doi.org/10.1016/0959-1524\(96\)00009-1](http://dx.doi.org/10.1016/0959-1524(96)00009-1)
- [9] Li, K., Guo, Y., Lavery, D., He, H. and Fei, M. (2013) Distributed Adaptive Learning Framework for Wide Area Monitoring of Power Systems Integrated with Distributed Generations. *Energy and Power Engineering*, **5**, 962. <http://dx.doi.org/10.4236/epe.2013.54B185>
- [10] Lavery, D.M., Morrow, D.J., Best, R. and Crossley, P.A. (2009) Performance of Phasor Measurement Units for Wide Area Real-Time Control. *IEEE of Power & Energy Society General Meeting*, 1-5. <http://dx.doi.org/10.1109/PES.2009.5275578>
- [11] Lavery, D.M., Morrow, D.J., McKinley, A. and Cregan, M. (2011) Openpmu: Open Source Platform for Synchronphasor Applications and Research. *IEEE of Power and Energy Society General Meeting*, 1-6. <http://dx.doi.org/10.1109/PES.2011.6039607>
- [12] Lavery, D.M., Best, R.J., Brogan, P., Khatib, I.A. and Vanfretti, L. (2013) The Openpmu Platform for Open-Source Phasor Measurements. *IEEE Transactions on Instrumentation and Measurement*, **62**, 701-709. <http://dx.doi.org/10.1109/TIM.2013.2240920>
- [13] Balabin, M., Go'lrner, K., Li, Y., Naumkin, I. and Rehtanz, C. (2010) Evaluation of Pmu Performance during Transients. 2010 *IEEE International Conference on Power System Technology (POWERCON)*, 1-8.
- [14] Ku, W., Storer, R.H. and Georgakis, C. (1995) Disturbance Detection and Isolation by Dynamic Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, **30**, 179-196. [http://dx.doi.org/10.1016/0169-7439\(95\)00076-3](http://dx.doi.org/10.1016/0169-7439(95)00076-3)
- [15] Dunia, R. and Joe Qin, S. (1998) Subspace Approach to Multidimensional Fault Identification and Reconstruction. *AIChE Journal*, **44**, 1813-1831. <http://dx.doi.org/10.1002/aic.690440812>