

# Switching Markov Gaussian Models for Dynamic Power System Inertia Estimation

Xue Cao, Bruce Stephen, *Senior Member IEEE*, Ibrahim Abdulhadi, *Member IEEE*, Campbell Booth, Graeme Burt, *Member IEEE*

**Abstract**— Future power systems could benefit considerably from having a continuous real-time estimate of system inertia. If realized, this could provide reference inputs to proactive control and protection systems which could enhance not only system stability but also operational economics through, for example, more informed ancillary reserve planning using knowledge of prevailing system conditions and stability margins. Performing these predictions in real time is a significant challenge owing to the complex stochastic and temporal relationships between available measurements. This paper proposes a statistical model capable of estimating system inertia in real time through observed steady-state and relatively small frequency variations; it is trained to learn the features that inter-relate steady-state averaged frequency variations and system inertia, using historical system data demonstrated over two consecutive years. The proposed algorithm is formulated as a Gaussian Mixture Model with temporal dependence encoded as Markov chains. Applied to the UK power system, it produces an optimized mean squared error within  $0.1 \text{ s}^2$  for 95% of the daily estimation if being calibrated on a half-hourly basis and maintains robustness through measurement interruptions of up to a period of three hours.

**Index Terms**— Real-time Inertia Estimation, Gaussian Mixture Models, Markov Chain, Model Training.

## I. INTRODUCTION

Due to the increasing penetration of renewable generation and an associated need for more flexible loads and technologies such as energy storage and demand response, the prevailing power network will inevitably exhibit less predictable and more pronounced dynamic behavior in the future [1]. In addition to the potential network reinforcement required, one of the wider implications is the absence of an inherent inertial frequency response from renewable sources, unlike large synchronous machines which naturally provide a valuable damping effect to any system frequency excursion based upon their inertia and amount of stored energy in their rotating masses. The absence of inertia from renewable sources will act to reduce overall system inertia and could compromise future system stability. Consequently, accelerated and magnified frequency oscillations and significant power flow volatility could occur more frequently in comparison with the behavior of present-day power networks [1]. Under such conditions, faster control action will be required or there may be potential for mal-operation of protection systems, such as RoCoF based loss-of-main protection and under-frequency load shedding (UFLS) [2, 3].

The continuous awareness of power system inertia, which would change more frequently in future due to the intermittent nature of renewable energy sources connected to the system, would be beneficial and could assist in identifying and quantifying system stability issues in real time. Furthermore, it offers a platform to pre-assess system frequency behavior in response to hypothetical disturbances. However, it is unrealistic to achieve a continuous online inertia estimation using existing techniques, which are either limited to post-event analysis using the Swing Equation [4, 5] or are implemented using an equivalent system model, which requires a large quantity of system measurements in practice [6, 7].

In order to achieve real-time inertia estimation while addressing the above issues, this paper formulates a Switching Gaussian Markov Model (SMGM) from the observed variations of frequency and inertia in historical measurements. SMGM accommodates the complex and multiple forms of dependencies inherent in the relationship between frequency and inertia variations by encoding a switching regime using a finite mixture model [8]. Moreover, the introduction of Markov dynamics provides a more accurate dependency structure by capturing temporal context from the time-series observation data.

The main contributions of the paper and reported method are as follows:

- 1) Due to the complex stochastic and specific temporal dependencies between system frequency and inertia, a novel multivariate mixture modeling approach to characterizing these dependencies has been proposed based on a combination of non-Gaussian distribution and Markov dynamics which follows a time-series skip-k transition [9].
- 2) The continuous estimation of power system inertia is achieved online taking only steady-state system frequency as input, periodically calibrated using generation dispatch information, attributing to the flexibility and accuracy of SMGM which encodes a switching regime and tracks the mixed-order Markov dynamics of a latent variable that is indicative of a particular frequency-inertia variation relationship.

The paper is organized as follows: Section II reviews the state-of-art methodologies of inertia estimation. Through analyzing the relationship between system frequency and inertia, Section III identifies the dependency structure for inertia estimation. In section IV, the constituent elements required to formulate the model are presented, specifically the use of a Gaussian Mixture Model (GMM) and the encoding of Markov dynamics. Section V presents issues associated with the accuracy of estimation and the influence of selection of parameters/models. Validation of the

proposed model and discussion of performance and implementation issues are covered in section VI along with an illustration of the robustness of the model when input measurements are not available for periods of time. Section VII contains concluding remarks and explores issues and future work associated with practical deployment of the technique on an actual network.

## II. EXISTING METHODS FOR INERTIA ESTIMATION

Existing inertia estimation methodologies can be classified as using transient frequency based Swing Equation, deformed Swing Equation and recursive state-based estimation.

### A. Transient frequency based inertia estimation

The Swing Equation is widely used in transient frequency based inertia estimation. It describes the dynamic change in rotor angle ( $\delta$ ) when a disturbance on the network upsets the balance between generator mechanical power ( $P_m$ ) and electrical power ( $P_e$ ) [10]. This phenomenon is represented in Equation (1), where  $H$  is the inertia constant of the generation unit and  $\omega_s$  is the rotor speed.

$$\frac{2H d^2\delta}{\omega_s dt^2} = P_m - P_e \quad (1)$$

Thresholds for identification of events and the suitability of applying such a polynomial approximation are introduced in [4, 11], along with a 500ms sampling window for improving the accuracy of the estimation. The location of phasor measurement units (PMUs) and the system loading conditions are shown to influence the accuracy of the estimation. In contrast to [11], [12] proposes a moving average filter used for inertia estimation by comparing the filtered active power and the derivative of frequency stored over a certain time period. Improvements are made in [5] by introducing consecutive sum from which the time of the disturbance and the inertia at that time can be derived. However, these methods are all limited in that the onset of a disturbance is required to determine system inertia. Moreover, the subsequent time needed for inertia computation delays effective actions necessary to reduce the impact of disturbances on system stability (e.g. through pre-emptive control actions). Furthermore, their applicability to future systems may be negatively influenced from the accuracy of measurement perspective, as a result of more dynamic (small signal and transient) behavior due to reduced inertia levels.

### B. Deformed swing equations

In this approach, the Swing Equation is utilized to formulate the equations of inertia estimation with specific system measurements (i.e. voltage, current, phase angle).

A fifth order polynomial approximation equation is applied in [13] to fit the rate of change of frequency curve where the estimated coefficients of the polynomial are used for inertia computation. In [14], a system inertia estimation is obtained from an expression of bus voltage and oscillatory frequency. However, this is only applicable to radial networks, while mesh networks are invariably used in practical power systems. A single machine infinite bus model is adopted in [7]. Inertia of an individual machine is

derived from the division of the change in electrical power by the third derivative of the rotor angle. However, this method is limited when extended to a multi-machine system.

### C. Recursive state estimation

According to real-time selected observations of system topology and measurements (e.g. voltage, phase angle), the states of a statistical model and its parameters, which equivalently represent the investigated network model, are updated in a recursive manner to provide an estimate of prevailing system inertia.

The update of system inertia in a multi-machine system is achieved in [6] through the iterative updates based on Least-squares Method and Newton-Raphson methods. However, the dynamic states of each generation model need to be known for deriving the sensitivity matrix and analyzing eigenvalues to provide an estimate of inertia. In practice, all generation plant (i.e. large scale plus DG and micro-generation) is rarely monitored in the control center and not all generation units in the system are responsible for inertial frequency response, meaning this method could under/over-estimate system overall inertia by missing or wrongly accumulating contributions from some of the generation units. An Extended Kalman Filter is applied in [15, 16] to calibrate system inertia as a prediction correction process by minimizing mean squared error (MSE) between predicted and actual values of inertia. MSE is a risk function that reflects the difference between actual observations and estimates, expressed as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (2)$$

$n$  is the total number of variables,  $\hat{y}_i$  is the estimate and  $y_i$  is the true value.

Of course, in practice, actual inertia measurements are not available; limiting the practical applicability of this approach.

## III. DEPENDENCY STRUCTURE FOR FREQUENCY AND INERTIA

### A. Non-Linearity and Non-Stationarity Characteristics

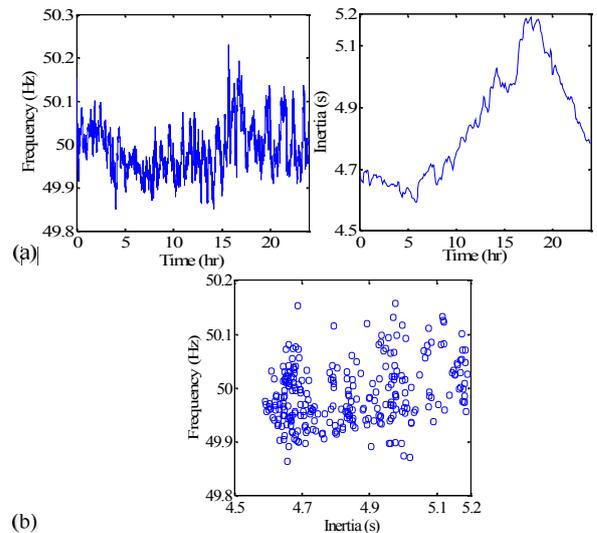


Fig. 1. Typical mapping of system frequency and inertia variations over the course of a day in: (a) time series, (b) scatter plot.

Typical daily frequency and inertia variations from the UK grid are shown in Fig. 1. The frequency data is recorded

from a PMU installed at the University of Strathclyde with a 20 ms reporting rate [17]. The inertia data is derived from historical generation dispatch data from [18] with each generation source assigned with a typical inertia value based on its type [19, 20]. It is computed as depicted in Equation (3) using the product of the sum of all individual inertias ( $H_n$ ) and the installed capacity of each individual generation unit ( $S_n$ ) over the total installed generation capacity. It is assumed that all the generators are working under the same power factor and efficiency, therefore, given the same value of power output ( $P_n$ ) and its rating  $S_n$ . The subscript ' $t$ ' corresponds to each generator. In contrast to the frequency data, the historical dispatch information is only available at 5 minute intervals due to limited access to commercial data.

$$H_{eq} = \frac{\sum_{i=1}^n H_n * S_n}{\sum S_n} \quad (3)$$

As shown in Fig. 1, the daily trajectories of frequency and corresponding inertia can be considered as non-linear in their temporal characteristics and non-stationary in terms of their range of observed values. These characteristics can be attributed to variations in system demand over time, the mix of generation on the system and response of generation control systems. Such system dynamics cannot be learnt or reflected through Equation (3).

Rather than using a single distribution to represent the hidden relationship between frequency and inertia, a complex set of inter-dependencies and relationships resulting from the coupling of non-linear and non-stationary behaviors necessitates a more representative model that goes beyond the linear Gaussian dynamics of traditional regression models.

### B. Temporal dependency analysis

The correlation coefficients linking frequency and inertia (including derived variables, e.g. rate of change of frequency) have been examined to establish the best pair of variables that reflects a strong dependency between frequency and inertia. This is achieved over 5 minute intervals where average frequency variation values are used between each sampling period. The data sets used were extracted from the available recordings of PMU data at the University of Strathclyde, University of Manchester and Imperial College London which form part of a UK monitoring network [17]. 100 days were randomly selected from years 2013 to 2015, giving 28,800 training samples which encompass daily and seasonal factors that may affect the analyzed variables. It should be noted that these data was checked to ensure they were taken during normal operating conditions ( $\pm 2\%$  of nominal frequency).

Fig. 2 presents statistics for the four (of a possible 25) stronger dependent variable combinations; ' $\Delta H/\Delta f$ ' is the variation of frequency/inertia between samples; ' $(\Delta H/\Delta t)/(\Delta f/\Delta t)$ ' is the derivative of change of frequency/inertia; ' $var$ ' introduces a moving window for filtering the variances of the inertia/frequency variation.

In practice, it is hard to derive inertia values from the estimated second derivatives of inertia accurately owing to the inherent cumulative error in estimating higher order derivatives from lower order derivative estimates. Meanwhile, the averaged mean of the correlation coefficient

between  $\Delta f$  and  $\Delta H$  is higher than other groups which generally vary between  $-0.3$  to  $0.3$ . Therefore,  $\Delta H$  and  $\Delta f$  are established as the two dependent variables as being the relatively strongest correlation pair. That is the reason behind the use of the probability model - it accommodates the noise associated with weaker relations. The computation of current system inertia is defined as the sum of estimated variation of system inertia and the inertia value derived at the latest known state:

$$H_t = H_0 + \sum_{\tau=1}^{\tau=k} \Delta H_{est,t|\Delta f} \quad (4)$$

where  $H_0$  is the last observed system inertia;  $\Delta H_{est,t|\Delta f}$  is the estimated variation of system inertia under an observed system frequency variation  $\Delta f$ ;  $H_t$  is the estimated system inertia at the time  $t$  over  $k$  accumulated periods.

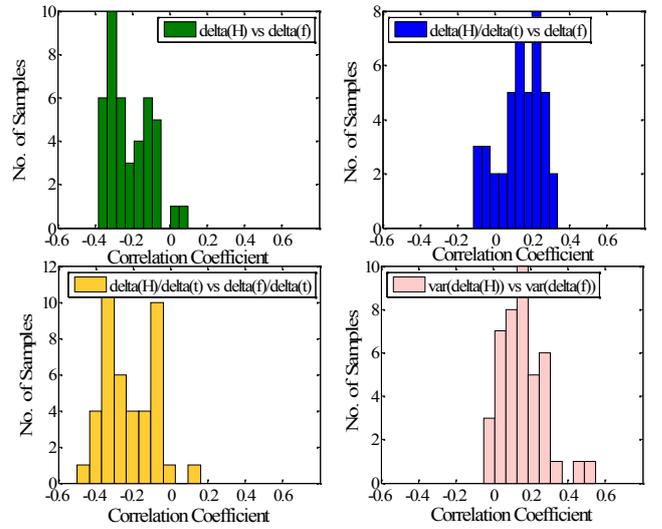


Fig. 2. Correlation coefficient of four relatively stronger combinations of frequency and inertia derived variables

## IV. OFFLINE TRAINING ALGORITHM FOR THE SWITCHING MARKOV GAUSSIAN MODELS

As previously noted, generation mix and control system operation are variables that will influence the inertia/frequency relationship - these are not typically measured and/or available and therefore can only exist in a model as latent or hidden variables. This concurs with the suggestion in Fig. 1 that no single distribution fits the joint distribution of frequency and inertia; observation data rarely fits a parametric distribution of a known form, given the potential for higher order statistics such as skewness and kurtosis being exhibited [21]. A flexible means of representing probability distributions of such an arbitrarily complex form is to use a finite mixture model - this comprises a set of simple parametric distribution, linearly combined and weighted to best fit the implied distribution of the observed data [22, 23]. A mixture of  $M$  Gaussians would be used to express the probability of observation  $\mathbf{x}$  of dimension  $d$  as:

$$P(\mathbf{x}) = \sum_{m=1}^M P(m) \frac{1}{2\pi \left| \Sigma_m \right|^{\frac{d}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu_m)^T \Sigma_m^{-1}(\mathbf{x}-\mu_m)} \quad (5)$$

$\mathbf{x}$  is the observation variable, in this case is the frequency variation given by a specific inertia variation,  $P(m)$  is the

mixing weight or probability of the mixture component  $\pi$  occurring and  $\mu$  and  $\Sigma$  are the Gaussian mean and covariance respectively of each mixture component. Matrices and vectors are shown bold.

Fig. 3 depicts a five-step process to train the mixture model offline for a representative expression of the dependency between frequency and inertia variations.

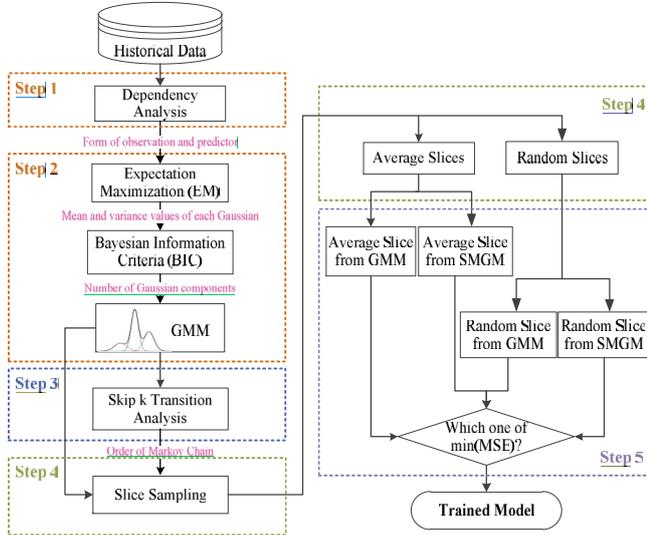


Fig. 3. Offline training of the SMGM

### 1) Step1: Dependency analysis

Training of the model starts with exploring the dependency pair where the mixture model can be formulated on. This has been demonstrated in Section III. B. where the underlying frequency-inertia dependency has been expressed as Equation (4).

### 2) Step2: Selection of components of GMM.

An iterative process is executed to select the number of mixtures and parameters of each Gaussian which best profile the joint distribution of frequency and inertia variations. Expectation Maximization (EM) [24, 25], which maximizes the probability of fitting observation data with candidate models, is a commonly used approach to estimate the parameters of an individual distribution component as well as the values of the mixing proportions. This is followed by the probability calculation running under different components of Gaussians. The one producing the minimum Bayesian Information Criterion (BIC) [26, 27] value is selected as the best number of mixtures. GMM is then summed from each individual Gaussian as a proportion of its corresponding weight factor which is assigned to identify the contribution from an individual Gaussian.

### 3) Step 3: Order Selection of the SMGM.

The mixed-order Markov Chain is introduced in Step 3 as a hypothesis to improve estimation accuracy. It preserves a temporal dependency where current state will be modeled in response to the previous one or more (possibly temporally discontinuous) states in a non-deterministic way [28, 29]. By assuming multiple underlying operating regimes, the mixed-order Markov Chain is designated to accommodate the complex stochastic and dependency structures inherent in the system. Such a temporal context is expected to express the relation between variations of frequency and inertia in greater detail. In order to determine which prior states are regressive with the current state, the “skip k

transition autoregression” technique [30, 31] is employed where k means the irrelevant states in between the current state and prior regressive states.

### 4) Step 4: Slice Sampling

The fitted SMGM now takes the form of a mixture distribution that represents variations in inertia that are conditional on observed variations in frequency in present and several past states. Predicting inertia variations, therefore, necessitates sampling from this mixture distribution which in turn, presents a challenge of dealing with non-Gaussian distributions. To avoid loss of generality, slice sampling is used to draw samples from the mixture distribution [32]. This allows distributions of an arbitrary form to be sampled from, without any specialization of the procedure. However, due to the fact that samples are randomly generated within the given range, there is no justification for selecting any individual sample. Therefore, comparisons will be made between the approaches of averaging through the entire set of slice samples and randomly selecting from the samples.

### 5) Step 5: Comparison of MSEs.

Optimal model selection is achieved by analyzing the MSEs from a total of four combinations of slice sampling techniques and with/without Markov Chain. The model with minimum MSE is selected as the final inertia estimation model. By incorporating both the variance of the estimate and its bias, MSE is superior to techniques like sum squared error therefore, made it the selection criterion in this research.

## V. IMPLEMENTATION FOR SYSTEM INERTIA ESTIMATION ONLINE

In order to examine the validity of proposed method, UK power system is conducted to test the performance of the previously formulated SMGM on real-time inertia estimation. The most representative SMGM model for UK system is trained as below with the same groups of data set for frequency-inertia dependency analysis in Section III. B. The frequency data which exceeds the  $\pm 2\%$  operational limits was not taken in the training data set.

### A. Selection of Gaussian Mixtures

In Section IV, the dependency of the GMM cardinality on the components of the mixture was articulated through the BIC expression for the model. Comparison is now made across three Gaussian covariance structures: spherical (independent variables with a common variance), diagonal (independent variables) and full (linear dependence) as shown in Table I.

It should be stated that the training of the GMM is performed offline and the precision and computation efficiency are governed by the choice of particular covariance structure. Thereby, the computational efficiency is not as important compared to the precision of capturing more observation information and the complexity of mixture components reflected by the number of mixtures. Based on the historical training data set, only 20 Gaussians are needed if using the full covariance matrix, which is less complex and represents observations in a more comprehensive way (shown under ‘Covariance Structure’ in TABLE I). To this end, the 20-component mixture with full covariance matrix

is selected for best describing the dependency relation.

An example of 20-component Gaussian mixture model is shown in Fig. 4 based on the conditional distribution of historical inertia variations on frequency variations. The probability of each Gaussian given an inertia variation is labeled individually in the 2-D scatter plot (Fig. 4 (a)) in which each color represents one Gaussian distribution. The contribution of each Gaussian component to the mixture model is shown in Fig. 4 (b) as individual weight factors.

TABLE I  
COMPARISON BETWEEN VARIOUS LINEAR GAUSSIAN  
DEPENDENCY STRUCTURES

|                      | Spherical   | Diagonal  | Full  |
|----------------------|---|---|---|
| Precision            | low   | medium  | high  |
| Computation          | high  | medium  | low   |
| Covariance Structure |  |  |  |
| Number of Components | >50   | 26  | 20  |

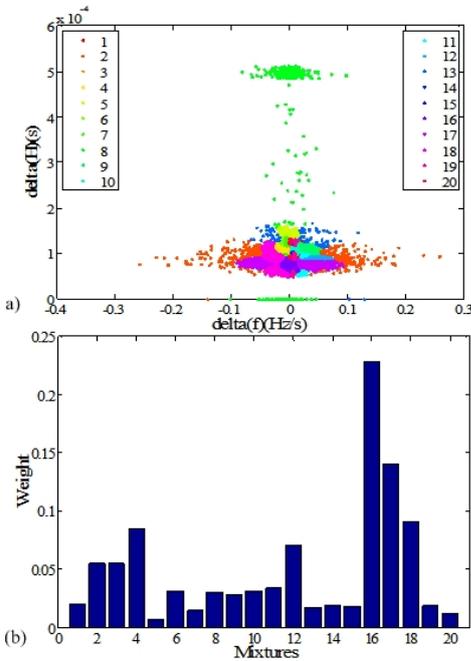


Fig. 4. (a) Labeling of 20 mixture components, (b) weight factors, from training data set of inertia variations conditional on frequency variations

TABLE II  
COMPARISON OF DIFFERENT IMPLEMENTATIONS OF SMGM

|                             | Zero-order SMGM | Fourth-order SMGM                         |
|-----------------------------|-----------------|---|
| No. of Gaussian Components  | 20              | 13  |
| Covariance Matrix Dimension | Full            | Diagonal                                  |
| Skip Transition State       | $t_k$           | $t_k, t_{k-2}, t_{k-3}, t_{k-4}, t_{k-5}$ |

### B. Switching Markov Model

Further reduction of estimation error has been assessed by the introduction of a switching Markov chain. In this case, a ‘Stepwise Regression’ technique is employed to achieve the ‘Skip k transition’ for selecting the dependent-states. It allows the action of automatically entering/removing

observers one by one in a justifiable and reasonable manner [33]. The temporal analysis given in Table II returns two forms of SMGM – zero-order which is just the original GMM independent of the time series and a fourth-order SMGM. In this case, the fourth-order SMGM denotes the prior states which have dependency with current state start from the second state till the fifth state in the past. Those relevant states are subscripted as  $t_{k-2}, t_{k-3}, t_{k-4}, t_{k-5}$ .

### C. Prediction of Inertia Variation using Slice Sampling

As stated earlier, the final estimates will be taken from the sampled slices from such a conditional distribution to avoid loss of the generalisation. However, the larger sample size will give more accurate indication while takes longer time to compute. As shown in Fig. 5, a size of 1000 slices is sampled as the trade-off between accuracy (determined by MSE) and efficiency. A local minimum of MSE is reached when the number of sampled slices equals to 1000 with a significant reduction of around 5% for samples ranging from 50 to 2000.

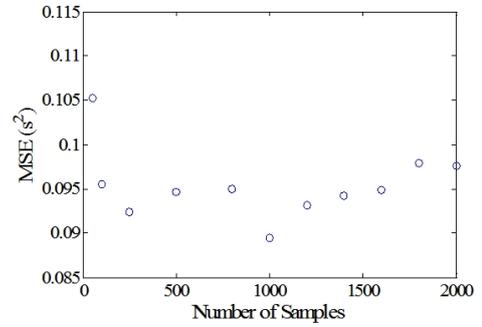


Fig. 5. Impact of slice sampling size on the accuracy of estimation

Thus far, the procedure for training the GMM has been described. The following sections focus on verifying the hypothesis that the encoding of Markov dynamics will improve the estimation accuracy.

### D. Online OCC (OCC)

To minimize the accumulated error in Equation (4) from summing the generated inertia variation estimate each time to derive present system inertia, the concept of OCC is introduced. It refers to a certain period of time after which the estimated inertia value will be corrected according to the system measurements. In this paper, the generation dispatch data in UK system is employed for calibration against an equivalent reference inertia value derived from Equation (3). Such information is available every thirty minutes on [18] which gives the minimum OCC of half an hour.

### E. Performance of SMGMs

As shown in Fig. 3, a total of four SMGM models need to be compared and the one of minimum MSE will be selected as the most representative UK model. The maximum and averaged MSEs from the 1000 samples are recorded in Table III, in association with various durations of OCC.

In the OCC perspective, the following characteristics are summarized:

- Both the maximum and averaged MSEs increase with increasing OCC times, regardless of the selection manner and orders of SMGM. A half-hourly OCC

should be chosen to provide the best estimation accuracy. Such data is readily available from [34].

- Three hours is the OCC boundary, above which errors in estimation grow significantly. For an OCC shorter than three hours, the average MSE remains below  $0.1s^2$ , even adopting the random slice selection approach.
- Choosing randomly from the sliced samples gives a around higher error (63% for max MSEs and 90% for averaged MSEs) in contrast to the average slice values for both SMGMs and a local minimum of 49.5% is experienced when the length of OCC reaches to three hours.
- The fourth-order SMGM can successfully improve estimation accuracy by reducing mean MSE of 8.8% for the random sample selection approach and 11.8% for the average sample selection approach. The greatest improvement by Markov dynamics is reached at a one-hour OCC where the maximum MSE is reduced by 17.6% and the averaged MSE is diminished by 8.7%.

TABLE III  
MAX AND AVERAGED MSEs FOR DIFFERENT MODELS

| OCC (hour) | MSE for Zero-order SMGM ( $s^2$ ) |          | MSE for Fourth-order SMGM ( $s^2$ ) |          |
|------------|-----------------------------------|----------|-------------------------------------|----------|
|            | Random Slice                      |          |                                     |          |
|            | Max                               | Averaged | Max                                 | Averaged |
| 0.5        | 0.1372                            | 0.0632   | 0.1285                              | 0.0597   |
| 1          | 0.2103                            | 0.0823   | 0.1732                              | 0.0751   |
| 3          | 0.6291                            | 0.4143   | 0.6999                              | 0.3379   |
| 12         | 1.3640                            | 0.6493   | 1.1320                              | 0.5123   |
| 24         | 2.5148                            | 0.9857   | 2.1953                              | 0.9132   |
|            | Averaged Slice                    |          |                                     |          |
|            | Max                               | Averaged | Max                                 | Averaged |
| 0.5        | 0.0601                            | 0.0023   | 0.0588                              | 0.0020   |
| 1          | 0.0635                            | 0.0071   | 0.0604                              | 0.0039   |
| 3          | 0.3175                            | 0.0298   | 0.3072                              | 0.0243   |
| 12         | 0.4001                            | 0.0901   | 0.3652                              | 0.0779   |
| 24         | 0.6989                            | 0.1703   | 0.6810                              | 0.1600   |

The above analysis shows that the average sample selection approach and the encoding of Markov Chain are two effective means to reduce the estimation errors. However, the performances using averaged slice sample could fail to trace the features of inertia variation as a result of the over-smoothing of individual variations. This deficiency is reflected in Fig. 6 from the comparison between referenced inertia trajectories from generation dispatch and corresponding model predictions using SMGMs over the course of one day (in the training data set). Missing of the variation features is more obvious with the increase of OCC through comparing the performance presented in Fig. 6 (c) and (d) on a three-hour basis with the results shown in Fig. 6 (a) and (b), where the OCC is 30 minutes. Zoom-in performance for a period of three hours in Fig. 6 (b) and (d) are depicted in Fig. 7. The feasibility of the SMGM is shown in the tracking of the system inertia variations and the changing system trends. The proposed model, if drawing estimation sample randomly from the conditional distribution, is shown higher accuracy in tracking these changes. Besides, a relatively large error appears with an increase in the period of the OCC.

To summarize, of the four investigated methods for

improving estimation accuracy, the fourth-order SMGM, which adopts a random selection approach from sampled 1000 slices and being calibrated half-hourly, delivers the most accurate estimation and tracking performance.

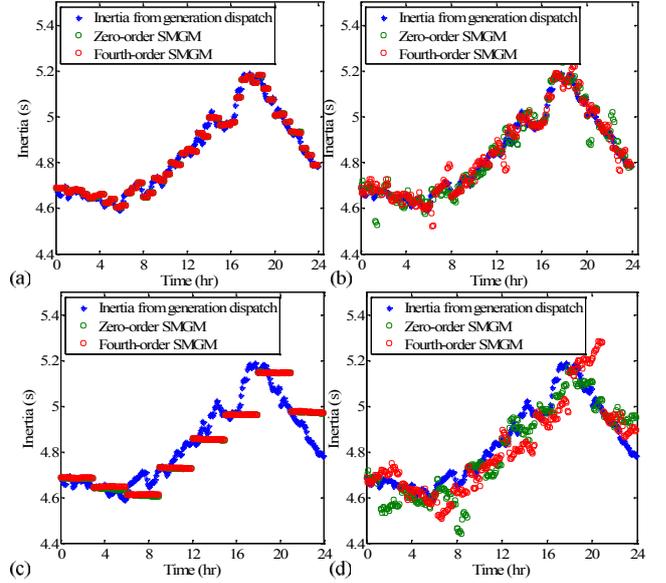


Fig. 6. Performance of SMGM algorithm with (a) half-hourly OCC and averaged sample selection; (b) half-hourly OCC and random sample selection; (c) 3-hour OCC and averaged sample selection; (d) 3-hour OCC and random sample selection.

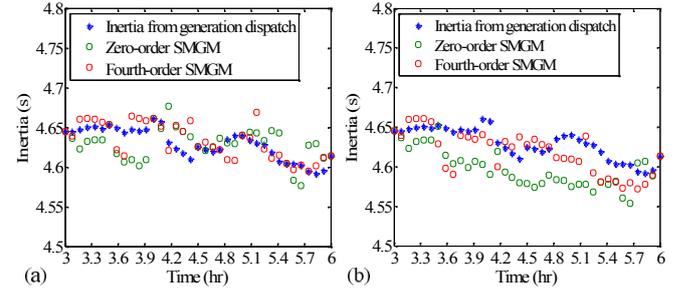


Fig. 7. An example of zoom-in figure of SMGM performance from (a) Fig. 6. (b) and (b) Fig. 6. (d)

## VI. VALIDATION AND DISCUSSION

### A. Model Validation

The online estimation performance of the formulated SMGM for UK system is validated in this section. The UK power system is used in each validation approach, utilizing frequency data recorded from PMU data recorded at the authors' laboratory (part of a network of University deployed PMUs [17], and historical generation dispatch schedule [18].

#### 1) Estimating post-event system inertia for recorded system disturbance

Fig. 8 shows the estimation of system inertia for Apr. 19th 2011 with a zoom-in after loss of a large generating plant in the north at around 6:36 UTC. Both of the SMGMs are assigned to half-an-hour OCC and it limits 92% of the maximum MSEs within  $0.005 s^2$ . More importantly, the zoom-in figure, which shows a general estimation error lower than 1%, indicates the ability of the proposed SMGM to accurately estimate the drop in system inertia post-disturbance.

## 2) Estimating pre-event system inertia for recorded system disturbance

The data used relates to the loss of the an interconnector on Sep. 30<sup>th</sup> 2012 at around 14:03 UTC [35]. The system inertia calculated using transient system data by Equation (1) is 4.659s. The estimated system inertias using the proposed algorithm are 4.427 s for zero-order SMGM and 4.439 s for fourth-order SMGM as shown in Fig. 9. The system inertia level theoretically should stay the same prior to and just after the event. Therefore, the inertia value estimated pre-disturbance using the proposed model can be compared with the inertia value calculated from the Swing Equation. Since the errors of both SMGMs (0.054 s for zero-order SMGM and 0.047 s for fourth-order SMGM) lie within the threshold in Table III, the proposed model is verified as being able to estimate system inertia prior to a system event and thus could provide useful input data to proactive control and protection functions.

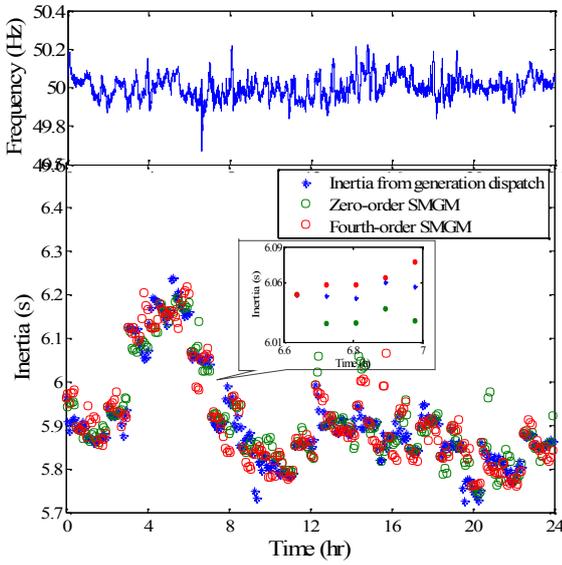


Fig. 8. Inertia estimation for a generator loss

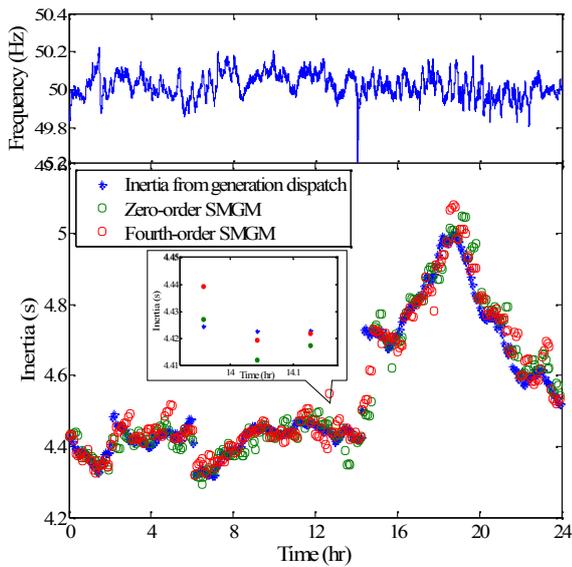


Fig. 9. Inertia estimation for the loss of an interconnector

## B. Discussion on estimation accuracy

The proposed algorithm has a number of parameters that must be selected properly for a more comprehensive representation of the dependency structure; this has

implications for the accuracy of real-time inertia estimation using the trained model. The parameters which have the greatest impact on the performance of estimation accuracy are listed below:

- Types of conditional distribution covariance matrix
- Number of components in the Gaussian mixture
- Size of the sampling slice sampler
- Approach to select predicted value from samples
- Order of skip-k transition SMGM
- Length of the OCC of the model

The covariance matrix, which can be specified to accommodate various dependencies, from spherical, to diagonal, to full, increases the precision of representing the observation but must be considered in association with the complexity of the model as a potential of increasing mixture components. The number of slice has little impact on obtaining the estimated value but a slight increase is desirable for an increased accuracy. Selecting sliced value from a random approach is better as the averaged value loses the details of changing trend of frequency that varies over time. Moreover, the increase in OCC will introduce more errors in inertia estimates due to the accumulated error in every subsequent computation. Finally, the selection of the model parameters needs taking into account the practical application's accuracy requirements.

## C. Inertia estimation in practice

Equation (4) states that the proposed method estimates current system inertia by summing the inertia value at a previous state with its estimated variation given by frequency variation. There will inevitably be an error accumulating over time as the estimated current system inertia is based on a number of aggregated estimations in the past.

Besides, the actual system inertia, which in this study is primarily calculated from generation dispatch data relating to large generators on the system, could be under-estimated as there may be further inertia contributed from smaller generators, very small DG and motor loads. Furthermore, the unknown dynamics of system inertia within the 5-minute interval could also lead to estimation errors if an exceptional generation/demand change occurs.

For the application of such advanced techniques in the future, the aforementioned accuracy issues could be mitigated with training data of a higher time resolution and detailed system information, including the proportion of demand served by hidden generation at the distribution level.

## D. Robustness against loss of frequency data

To investigate the robustness of the proposed algorithm, a series of test scenarios for the loss of measured frequency data used by the inertia estimator are presented. In practice, the data loss can be resulted from the hardware/software. Test durations vary from 1 cycle (5 minutes) up to 60 cycles (5 hours). The input data (averaged frequency variation over every 5 minutes), during this period of time, is set to be either zero or keep the constant inertia value estimated just before detected input data loss.

Examples of maximum MSE are given in Fig. 10 using both SMGMs on a random selected date. With the increase of data loss cycles, an approximate linear increase of MSE

can be seen for SMGMs while an exponential character for constant inertia. For applying the criterion of keeping constant inertia value, the MSE is shown limited within  $0.05 \text{ s}^2$  which gives a more accurate estimation. Therefore, the fail-safe plan in deployment of SMGM in real-time inertia estimation is to maintain the last estimated inertia value until the successful delivery of input data. Moreover, the proposed model is considered to be robust against loss of frequency input up to three hours at which point an error of 10% for the actual system inertia level is obtained.

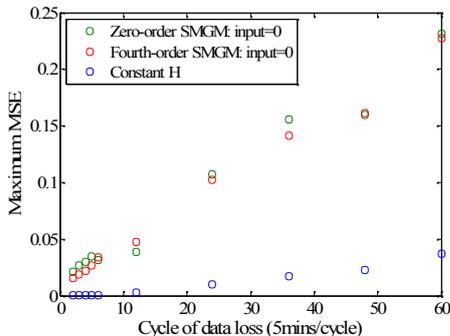


Fig. 10. Mean MSE for various durations of data loss on May 20th 2013

#### E. On the occurrence of large system disturbances

Due to the fact that the SMGM is trained with data taken during normal operating conditions (frequency varies within  $\pm 2\%$  of the nominal 50 Hz in UK system), the formulated SMGM then can only be valid to generate real-time inertia estimates during normal condition. If such boundary is exceeded, the inertia estimates will be invalid. It is a limitation of this methodology but it aims to provide valid inertia estimates to take actions before any disturbance could potentially occur. Meanwhile, such an inconsistency is temporary as the validation in Section VI.A has proven its feasibility to estimate system inertia post-disturbance.

#### F. Time resolution of implementation

The proposed model estimates system inertia on a 5-minute resolution basis which is governed but limited by the resolution of the available system data that has been used for model training. Higher resolution data sets could clearly provide enhanced accuracy, and the methodology remains flexible to accommodate future data provision. However it is the application which should dictate the estimation rate where quick and accurate estimation may be preferred for monitoring system status, while less frequent estimation may be preferred to limit adjustments in the case of adaptive control or protection applications.

With respect to the time resolution of inertia calibration, a half-hourly basis is recommended due to its lowest MSEs and it also satisfies the practical constraint where the shortest updating cycle of generation dispatch information is currently available every 30 minutes. However, if the data necessitated for calibration cannot be delivered every half an hour, a resolution of up to three hours is still acceptable for its monitoring.

#### G. Impact of PMU requisites

The use of data from PMU recordings in the authors' laboratory demonstrates the value of such frequency data to the application of SMGMs, and thus their value in inferring system conditions. Various requisites associated with PMU

deployment have at the same time been investigated.

This work is based on the use of the frequency variations averaged over three locations, from which a system-level insight is provided rather than a single PMU that could largely be affected by its local conditions. Moreover, it can also provide a consolidated justification to generate system overall inertia estimates which is also a system-level variable. Due to the fact that the system frequency variations are strongly correlated during normal operating conditions, impacts of the total number of PMUs and the exact locations of each PMU are not that significant. However, it is undoubted that the increasing number of PMUs could assist in providing a finer and more generic system vision. The noise and latency effects are minimized by the frequency data input to the SMGM being averaged (on a 20ms basis) over a 5-minute interval and also being averaged across different PMU locations.

However, it should be emphasized that the PMU is not the compulsory means of obtaining system frequency information but an accessible source from the authors' laboratory. Meanwhile, the practicalities of such and its role in implementing SMGM in future networks can be supported by the learning from projects such as EFCC [36] and VISOR [37].

#### H. Practical implementation of the SMGM algorithm

The generic nature of SMGM makes it readily applicable to various power systems for which it has been trained. Moreover, system changes, such as generation mix and the employment of electronic-based inverters, can be learnt through re-training the model with most up-to-date data.

Such flexibility and adaptability attribute to the employment of correlation, EM, BIC and stepwise regression. As previously described in this paper, they provide robust justifications for identifying the methods' parameters, such as the optimal numbers of Gaussian components and Markov orders. Widespread implementation of this approach is thus supported, whereby the SMGM is trained offline and then utilized online through exposure to frequency measurements and regular calibration with dispatch data.

Furthermore, the regular calibration points within a particular implementation (half-hourly rate was adopted in this paper) provide an opportunity to monitor inertia estimation errors. And if larger than expected errors emerged after a period of online use, then it may be appropriate to launch a new phase of offline training. Such behavior may be observed following changes in the power system such as further penetration of renewables or unit decommissioning. This feature may be integrated within an overarching application management system in order to maintain estimation accuracy over longer term use.

## VII. CONCLUSIONS

This paper proposed a novel approach to the modeling of frequency and inertia dependencies through which an estimation of system inertia can be obtained in real-time. The methodology is based on the optimization of the joint likelihood that the observed data follows a trained mixed-order Gaussian distribution. Unlike the methods using swing equation or other existing methodologies, this approach does

not rely on the occurrence of a disturbance to estimate inertia and can be applied online. More importantly, the hidden variables/states that exist but cannot be directly measured in dynamic system operations can be formulated over time.

Conducted with UK power system, a fourth-order SMGM is evaluated to deliver best inertia estimation performance. The proposed algorithm can estimate system inertia with an average MSE lower than  $0.1 \text{ s}^2$  on a half-an-hour calibrating basis, under an appropriate selection of covariance matrix, order of mixture model and the selection approach from samples during model training. Improvement of the estimation accuracy is achieved by introducing a higher Markov order to the Switching Gaussian model which establishes the dependency further from time series. The findings demonstrate the fourth-order skip transition SMGM adopting a random selection approach could deliver a more accurate estimation by reducing the maximum MSE up to 17.6% and the greatest improvement is observed at an OCC of one hour. Furthermore, the robustness of proposed model is also proved against measured frequency data loss of up to two hours for steady state conditions and capable to provide accurate inertia reference pre/post-disturbance.

The continuous inertia estimation model can be applied in various fields in future power networks where the variations of system inertia pose a significant stability issues. Such applications include early system diagnostics, frequency-related protection, regulation of power reserve and synthetic inertia control and management. Additionally, the model can be automatically re-trained and its performance (in terms of suitability and accuracy) could be continuously monitored by an overarching application management system.

Further improvements on the performance of proposed SMGM could also take into account other system variables, including loads which may consist of rotating mass (e.g. motors), demand side response (e.g. from electric vehicles), etc. Finally, the study of regional variations in system inertia should be undertaken with the assessment of PMU requisites (e.g. number, location) to realize this function.

## REFERENCES

- [1] N. Grid. *Electricity Ten Year Statement 2014*. Available: <http://www2.nationalgrid.com/UK/Industry-information/Future-of-Energy/Electricity-Ten-Year-Statement/>
- [2] F. Gonzalez-Longatt, E. Chikuni, and E. Rashayi, "Effects of the Synthetic Inertia from wind power on the total system inertia after a frequency disturbance," in *Industrial Technology (ICIT), 2013 IEEE International Conference on*, 2013, pp. 826-832.
- [3] C. Xue, I. Abdulhadi, A. Emhemed, C. Booth, and G. Burt, "Evaluation of the impact of variable system inertia on the performance of frequency based protection," in *Developments in Power System Protection (DPSP 2014), 12th IET International Conference on*, 2014, pp. 1-6.
- [4] D. P. Chassin, Z. Huang, M. K. Donnelly, C. Hassler, E. Ramirez, and C. Ray, "Estimation of WECC system inertia using observed frequency transients," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1190-1192, 2005.
- [5] P. Wall and V. Terzija, "Simultaneous Estimation of the Time of Disturbance and Inertia in Power Systems," *Power Delivery, IEEE Transactions on*, vol. PP, pp. 1-14, 2014.
- [6] G. Song and J. Bialek, "Synchronous machine inertia constants updating using Wide Area Measurements," in *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, 2012, pp. 1-7.
- [7] F. Lingling, M. Zhixin, and Y. Wehbe, "Application of Dynamic State and Parameter Estimation Techniques on Real-World Data," *Smart Grid, IEEE Transactions on*, vol. 4, pp. 1133-1141, 2013.
- [8] S. Frühwirth-Schnatter, *Finite mixture and Markov switching models*: New York : Springer, 2006.
- [9] G. J. McLachlan, *Finite mixture models*: New York : Wiley, 2000.
- [10] D. P. Kothari, *Modern Power System Analysis*: McGraw-Hill Education (India) Pvt Limited, 2011.
- [11] P. M. Ashton, G. A. Taylor, A. M. Carter, M. E. Bradley, and W. Hung, "Application of phasor measurement units to estimate power system inertial frequency response," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, 2013, pp. 1-5.
- [12] P. Wall, F. Gonzalez-Longatt, and V. Terzija, "Estimation of generator inertia available during a disturbance," in *Power and Energy Society General Meeting, 2012 IEEE*, 2012, pp. 1-8.
- [13] T. Inoue, H. Taniguchi, Y. Ikeguchi, and K. Yoshida, "Estimation of power system inertia constant and capacity of spinning-reserve support generators using measured frequency transients," *Power Systems, IEEE Transactions on*, vol. 12, pp. 136-143, 1997.
- [14] J. H. Chow, A. Chakraborty, L. Vanfretti, and M. Arcak, "Estimation of Radial Power System Transfer Path Dynamic Parameters Using Synchronized Phasor Data," *Power Systems, IEEE Transactions on*, vol. 23, pp. 564-571, 2008.
- [15] H. Zhenyu, D. Pengwei, D. N. Kosterev, and Y. Bo, "Application of extended Kalman filter techniques for dynamic model parameter calibration," in *Power & Energy Society General Meeting, 2009. PES '09. IEEE*, 2009, pp. 1-8.
- [16] K. Kalsi, S. Yannan, H. Zhenyu, D. Pengwei, R. Diao, K. K. Anderson, et al., "Calibrating multi-machine power system parameters with the extended Kalman filter," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1-8.
- [17] V. Terzija, P. Regulski, L. P. Kunjumammed, B. C. Pal, G. Burt, I. Abdulhadi, et al., "FlexNet wide area monitoring system," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1-7.
- [18] ELEXON. [Online]. Available: <https://www.elexonportal.co.uk/article/view/7324?cachebust=p5bbmvxyui>
- [19] P. Kundur, *Power System Stability and Control*: McGraw-Hill Inc., 1993.
- [20] F. c. s. a. p. m. p. b. w. g. B. Ekanayake, N. Jenkins, and G. Strbac. (2008) Frequency Response From Wind Turbines. *Wind Engineering*. 573-586. Available: <http://multi-science.metapress.com/content/82v333372668x067/fulltext.pdf>
- [21] *NIST/SEMATECH e-Handbook of Statistical Methods*. Available: <http://www.itl.nist.gov/div898/handbook/>
- [22] H. Xiaofei, C. Deng, S. Yuanlong, B. Hujun, and H. Jiawei, "Laplacian Regularized Gaussian Mixture Model for Data Clustering," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 23, pp. 1406-1418, 2011.
- [23] G. McLachlan and D. Peel, *Finite Mixture Models*: Wiley, 2004.
- [24] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B*, vol. 39, 1977.
- [25] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, 2nd Edition ed.: New York: Wiley-Interscience, 2008.
- [26] R. E. Kass and A. E. Raftery, "Bayes factors," *Journal of the american statistical association*, vol. 90, pp. 773-795, 1995.
- [27] K. Hirose, S. Kawano, S. Konishi, and M. Ichikawa4, "Bayesian Information Criterion and Selection of the Number of Factors in Factor Analysis Models," *Journal of Data Science*, vol. 9, pp. 261-270, April 2011.
- [28] L. Saul and F. Pereira, "Aggregate and mixed-order Markov models for statistical language processing," in *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing 1997*, pp. 81-89.
- [29] S. Fariaa and G. Soromenhob, "Fitting mixtures of linear regressions," *Journal of Statistical Computation and Simulation*, vol. 80, Feb. 2010.
- [30] D. v. Dijk, T. Teräsvirta, and P. H. Franses, "Smooth Transition Autoregressive Models — a Survey of Recent Developments," *Econometric Reviews*, vol. 21, pp. 1-47, 2002/01/04 2002.
- [31] A. Timmermann, "Moments of Markov switching models," *Journal of Econometrics*, vol. 96, pp. 75-111, 2000.
- [32] R. M. Neal, "Slice Sampling," *The Annals of Statistics*, vol. 31, pp. 705-767, 2003.
- [33] P. S. University. *Stepwise regression*. Available: <https://onlinecourses.science.psu.edu/stat501/node/88>
- [34] *neta*. Available: <http://www.bmreports.com/>
- [35] W. Hung, G. Ray, and G. Stein. (26 Oct. 2012). *Frequency Changes during Large Disturbances WG*. Available: <http://www.nationalgrid.com/NR/rdonlyres/D07952E0-2B58-426F-B61F-C77A239D3964/57270/Meeting1Presentation.pdf>
- [36] *Enhanced Frequency Control Capability (EFCC) Project*. Available: <http://www2.nationalgrid.com/Mediacentral/UK-Press->

releases/2014/%C2%A312-6million-boost-for-two-innovative-National-Grid-projects/

- [37] *Visualisation of Real Time System Dynamics using Enhanced Monitoring (VISOR) Project.* Available: <http://www.spenergynetworks.co.uk/pages/visor.asp>



**Xue Cao** is currently undertaking a PhD within the Institute for Energy and Environment at the University of Strathclyde. She received her BEng from both University of Strathclyde and North China Electric Power University. Her research interests include power system dynamic modeling, adaptive protection, system integrity protection and wide area protection.



**Dr Bruce Stephen (M'09, SM'14)** currently holds the post of Senior Research Fellow within the Institute for Energy and Environment at the University of Strathclyde. He received his B.Sc. from Glasgow University and M.Sc. and PhD degrees from the University of Strathclyde and is a Chartered Engineer. His research interests include Distributed Information

Systems, Machine Learning applications in Power System and Energy End Use.



**Dr Ibrahim F. Abdulhadi (S'08, M'11)** received the MEng degree at the University of Strathclyde in 2007. He is currently a Research Assistant within the university's Institute for Energy and Environment and obtained his PhD on adaptive power system protection in 2013. He has experience working for UK distribution and transmission network operators. His current research

interests include the application of PMUs for wide area protection, real-time power system simulation, hardware in the loop testing and communications applications in smart grids.



**Dr Campbell D. Booth** received the B.Eng. and Ph.D. degrees in electrical and electronic engineering from the University of Strathclyde, Glasgow, U.K, in 1991 and 1996, respectively. Currently, he is a Reader with the Department of Electronic and Electrical Engineering, Institute for Energy and Environment, University of Strathclyde. His research interests include

power system protection; plant condition monitoring and intelligent asset management; applications of intelligent system techniques to power system monitoring, protection, and control; knowledge management; and decision support systems.



**Prof Graeme Burt (M'95)** received his BEng in Electrical and Electronic Engineering and his PhD from the University of Strathclyde in 1988 and 1992 respectively. His research interests lie predominantly in the areas of power system modeling and simulation, power system protection, and experienced validation of advance control and protection.