

Measurement-based correlation approach for power system dynamic response estimation

Author

Bai, Feifei, Liu, Yong, Liu, Yilu, Sun, Kai, Bhatt, Navin, Del Rosso, Alberto, Farantatos, Evangelos, Wang, Xiaoru

Published

2015

Journal Title

IET Generation, Transmission & Distribution

Version

Accepted Manuscript (AM)

DOI

[10.1049/iet-gtd.2014.1013](https://doi.org/10.1049/iet-gtd.2014.1013)

Rights statement

© 2015 John Wiley & Sons Ltd. This is the peer reviewed version of the following article: Measurement-based correlation approach for power system dynamic response estimation, IET Generation, Transmission & Distribution, 2015, 9 (12), pp. 1474-1484, which has been published in final form at <https://doi.org/10.1049/iet-gtd.2014.1013>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

Downloaded from

<http://hdl.handle.net/10072/413500>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Measurement-Based Correlation Approach for Power System Dynamic Response Estimation

Feifei Bai^{1,2}, Yong Liu², Yilu Liu^{2,3}, Kai Sun², Navin Bhatt⁴, Alberto Del Rosso⁴, Evangelos Farantatos⁴ and Xiaoru Wang¹

¹*Southwest Jiaotong University, Chengdu, China*

²*University of Tennessee, Knoxville, USA*

³*Oak Ridge National Laboratory, Oak Ridge, USA*

⁴*Electric Power Research Institute, USA*

Abstract: Understanding power system dynamics is essential for online stability assessment and control applications. GPS-synchronized phasor measurement units (PMUs) and frequency disturbance recorders (FDRs) make power system dynamics visible and deliver an accurate picture of the overall operation condition to system operators. However, in the actual field implementations, some measurement data can be inaccessible for various reasons, e.g., most notably failure of communication. In this paper, a measurement-based approach is proposed to estimate the missing power system dynamics. Specifically, a correlation coefficient index (CCI) is proposed to describe the correlation relationship between different measurements. Then, the auto-regressive with exogenous input (ARX) identification model is employed to estimate the missing system dynamic response. The U.S. Eastern Interconnection (EI) is utilized in this paper as a case study. The robustness of the correlation approach is verified by a wide variety of case studies as well. Finally, the proposed correlation approach is applied to the real FDR data for power system dynamic response estimation. The results indicate that the correlation approach could help select better input locations and thus improve the response estimation accuracy.

1. Introduction

Presently, wide area measurement systems (WAMS) based on synchrophasors, e.g., phasor measurement units (PMUs)[1] and frequency disturbance recorders (FDRs) [2], can continuously provide high-sampling-rate, GPS-synchronized phasor measurements over wide areas, which enable the real-time monitoring of wide-area power system dynamics [3-4]. These measurement data contain very important information for power system researchers and operators to understand the power grid dynamics. Therefore, a lot of research has been carried out based on the real measurement data of the power grid, e.g., frequency prediction [5], inter-area oscillation analysis and damping control [6-7], and stability analysis [8]. However, there are still some issues, which need to be solved: firstly, PMUs/FDRs are usually installed in limited locations of the power grid, which makes it difficult to study the dynamic characteristics of certain locations of interest; secondly, some PMUs/FDRs may lose partial measurements due to the failure of communication between PMUs and the central operator, which may make the system unobservable; thirdly, some PMUs/FDRs may deliver bad data because of cyber attacks, which may cause wrong operation decisions [9-10]. Therefore, the dynamic information in these conditions should be accurately estimated for enhancing system stability and reliability.

Traditionally, missing data reconstruction methods are used to help find the missing measurements[11-12]. Gao et al. [11] formulated the missing PMU data reconstruction problem into a low-rank matrix completion problem, but the method only works for problems that are missing some data points. Traditional state estimators usually generate pseudo-measurements from historical data or context of the generation of pseudo-measurements to replace the missing data [12]. As for the most widely studied static state estimation, it could not capture the dynamics very well after disturbances [12-13]. Therefore, the dynamic state estimation is studied [14-17]. Aminifa et al. [14] proposed a dynamic state estimation approach to predict voltage magnitude and phase angle whenever an operation suddenly changed. Ghahremani and Kamwa [15] presented a dynamic state estimation process based on Kalman filtering techniques to simultaneously estimate synchronous generator rotor angle and speed. However, both of proposed methods depend on the system circuit model, which may not be adequately accurate for two reasons: on the one

hand, the power system circuit-based dynamic model in the simulation software could not include all the details of the power grid; on the other hand, the operation condition and topology change frequently (although usually not dramatically) in the power grid. Neither of these two aspects can be addressed by the existing circuit-based model. To avoid these shortcomings, measurement-based approaches had been proposed. Multi-layer feed forward artificial neural network (ANN) models were trained and used to estimate the generator rotor speed in a two machine system, whose estimation comparison results show that the measurement-based model is better than the classical circuit-based model [16]. Angel et al. [17] proposed transient rotor angle estimation using ANN in a five-generator system, the results show that the measurement-based method could be used in real time for the ultra-fast estimation. Besides, most of the previous dynamic state estimations have focused on small-sized systems, which did not involve lots of candidate input locations for the estimation model construction.

This paper proposes a measurement-based dynamic response estimation method, which could estimate a period of dynamic response, e.g., the entire frequency response after a generation trip. In addition, the validation of this approach is also carried out using FDR data from the U.S. Eastern Interconnection (EI). In contrast with all the methods above, the main advantages of the proposed approach are: (1) It is purely based on measurements, thus does not rely on the circuit model of the power grid. (2) The missing measurements are estimated by other strong correlated accessible measurements, rather than their own historical data. (3) The correlation between system dynamics can be shown in a visual way on a power grid map, which could be updated in real time.

The rest of the paper is organized as follows. Section 2 introduces the model structure used to estimate the missing measurements. The methodology to select the best input locations for the model construction is described in Section 3. Section 4 provides a case study of the proposed approach in the EI system. The robustness of the algorithm is demonstrated and strong correlation areas are defined in Section 5. The algorithm is applied to real measurement data for dynamic response estimation in Section 6. Conclusions are offered in Section 7.

2. Model Structure for Dynamic Response Estimation

System identification is the approach to model behaviors of a dynamic system based on measurement data [18]. An identification model could be constructed using the PMU/FDR measurements as model input (s) and output (s). However, it is unrealistic to develop a model that can include all the details of power system dynamic characteristics. Generally, the authors had to utilize a model that correctly captures the specific dynamic phenomenon or interaction for a particular study. Depending on the purpose of the study, the appropriate model of a given power system component could vary significantly. Therefore, a power system is usually modeled as a black box with finite measurement points, which is usually expressed in a transfer function form [18]. In this paper, the linear auto-regressive with exogenous input (ARX) model is used as a transfer function to estimate the system dynamics. The mathematical structure expression of the single-input single-output (SISO) ARX model structure is described by a linear difference equation [18]:

$$y(k) + a_1 y(k-1) + \dots + a_{n_a} y(k-n_a) = b_1 u(k-1) + \dots + b_{n_b} u(k-n_b) + e(k) \quad (1)$$

For simplification, (1) which can be rewritten as:

$$A(q) y(k) = B(q) u(k) + e(k) \quad (2)$$

where k is the data point index, and $e(k)$ is a white noise. n_a and n_b are the orders of the signal $y(k)$ and $u(k)$, respectively. $u(k)$ is the input signal and $y(k)$ is the output signal, which is the currently inaccessible system response, respectively. $A(q)$ and $B(q)$ are matrix polynomials in q^{-1} , meaning that $A(q)$ and $B(q)$ are matrices whose entries are polynomials in q^{-1} , as shown in (3) and (4), respectively. q^{-1} is a backward shift operator, which is the conventional definition of the z-transform [18].

$$A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \quad (3)$$

$$B(q) = b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \quad (4)$$

With the SISO ARX model structure in (2), the multi-input single-output (MISO) ARX model structure can be derived:

$$A(q) y(k) = \sum_{j=1}^{n_j} B_j(q) u_j(k) + e(k) \quad (5)$$

where j is the number of input signal, n_j is the order of the j^{th} input signal.

Because of the linear structure of the ARX model, the model parameters of a multi-variable ARX model can be estimated by a linear least-square (LS) estimation method which usually uses QR factorization for overdetermined linear equations to optimize the ARX model parameters and to minimize the function V_{LS} . The details of ARX structure and algorithm are discussed in [18-19].

$$\text{Min } J(V_{LS}) = \sum_{k=n_a+1}^N (\varepsilon(k))^2 \quad (6)$$

where N is the total number of the data points; n_s is order of the ARX model; the equation error $\varepsilon(k)$ is described by:

$$\varepsilon(k) = \hat{y}(k) - y(k) \quad (7)$$

where $\hat{y}(k)$, and $y(k)$ are the measured response and estimated response, respectively.

To evaluate the identified ARX model, a model fitness criterion is employed as the model accuracy index:

$$F = \left(1 - \frac{\sqrt{(\hat{Y} - Y)^2}}{\sqrt{(\hat{Y} - \bar{Y})^2}}\right) \times 100 \quad (8)$$

where Y , \hat{Y} , and \bar{Y} are the estimated response, measured response, and the mean value of the measured response, respectively. This index is used to reflect the accuracy of the model in describing system dynamics. A fitness of 100 means a perfect fit between the estimated response and the measured response, while a fitness of zero means the estimated response is no better than the mean value of the measured response.

For easier interpretation, a normalization process that converts the accuracy index from $(-\infty, 100]$ to $(0, 1]$ can be performed as

$$A = e^{(F/100)-1} \quad (9)$$

After obtaining the model orders and parameters, the identified models can be evaluated by comparing their outputs with the actual system responses. The objective function for the model evaluation is;

$$e^{(F/100)-1} > \varepsilon \quad (10)$$

where ε is the model accuracy threshold. ARX models with low accuracy index will be rejected while only the model with the acceptable accuracy can be selected as the final model for further estimation studies. The threshold of the accuracy index can be defined according to the study system and the event type [20].

3. Methodology of the Proposed Approach

A large power system indicates a great number of potential input locations. The number of input locations actually used by the model must be reduced to an acceptable level for the sake of model updating speed. Therefore, input location selection is a critical aspect of system identification since it directly affects the model accuracy and complexity [21-22].

In a power grid, the measurements in different locations have some underlying relationships, e.g., all the bus frequencies change similarly after a generation trip. In this paper, a linear correlation concept in signal processing is employed to describe the relationship between arbitrary measurement locations in power grids. Based on the correlation ranked results, the inputs of the measurement-based model will be selected. In [23], ARX model is used to estimate the system dynamics, but it ignores one important aspect—how to choose the input locations from hundreds of measured locations. The methodology of how to define the correlation between measurements is described as follows.

3.1 Correlation coefficient index

In power systems, a commonly used measurement-based coherency function [24-25] is defined as:

$$\gamma_{xy}(f) = \frac{|S_{xy}(f)|}{\sqrt{S_{xx}(f)S_{yy}(f)}} \quad |\gamma_{xy}| \leq 1 \quad (11)$$

where f is the frequency, γ_{xy} is the coherency between power system measured signals $\{x(t)\}$ and $\{y(t)\}$. $S_{xy}(f)$ is the cross-spectral density (CSD) function between $\{x(t)\}$ and $\{y(t)\}$, $S_{xx}(f)$ and $S_{yy}(f)$ are the power-spectral density (PSD) of $\{x(t)\}$ and $\{y(t)\}$, respectively. And these two signals are assumed as the wide-sense stationary random processes. This coherency function gives the linear correlation between two power system output signals as a function of the frequency. However, this frequency domain function is not convenient when dealing with a large number signals. On the other hand, the frequency domain function pays more attention to the coherency of the specific frequency point. The following equations will be used to derive the time domain correlation function for a wide frequency.

The cross-correlation function $R_{xy}(\tau)$, self-correlation functions $R_{xx}(\tau)$ and $R_{yy}(\tau)$ are given by the inverse Fourier transform of $S_{xy}(f)$, $S_{xx}(f)$ and $S_{yy}(f)$, respectively.

$$R_{xy}(\tau) = \int_{-\infty}^{+\infty} S_{xy}(f) e^{j2\pi f\tau} df \quad (12)$$

$$R_{xx}(\tau) = \int_{-\infty}^{+\infty} S_{xx}(f)e^{j2\pi f\tau} df \quad (13)$$

$$R_{yy}(\tau) = \int_{-\infty}^{+\infty} S_{yy}(f)e^{j2\pi f\tau} df \quad (14)$$

where τ is the time delay.

Applying the inverse Fourier transform to (11) and using (12), (13), and (14), the correlation function in the time domain is:

$$r_{xy}(\tau) = \int_{-\infty}^{+\infty} \gamma_{xy}(f)e^{j2\pi f\tau} df = \frac{R_{xy}(\tau)}{\sqrt{R_{xx}(\tau)R_{yy}(\tau)}} \quad (15)$$

The mathematical expectations of these two signals are $u_x = E\{x(t)\}$ and $u_y = E\{y(t)\}$, respectively. So the cross-correlation function $R_{xy}(\tau)$ and cross-covariance function $C_{xy}(\tau)$ are defined:

$$R_{xy}(\tau) = E\{x(t)y(t+\tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)y(t+\tau)dt \quad (16)$$

$$\begin{aligned} C_{xy}(\tau) &= E\left[\{x(t) - \mu_x\}\{y(t+\tau) - \mu_y\}\right] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \{x(t) - \mu_x\}\{y(t+\tau) - \mu_y\}dt \\ &= R_{xy}(\tau) - \mu_x\mu_y \end{aligned} \quad (17)$$

For the special case where $x(t) = y(t)$, the self-covariance function of $C_{xx}(\tau)$ and $C_{yy}(\tau)$ are:

$$C_{xx}(\tau) = R_{xx}(\tau) - \mu_x^2 \quad C_{yy}(\tau) = R_{yy}(\tau) - \mu_y^2 \quad (18)$$

If $\mu_x = 0$, $\mu_y = 0$ in (15), it can be obtained:

$$C_{xy}(\tau) = R_{xy}(\tau) \quad C_{xx}(\tau) = R_{xx}(\tau) \quad C_{yy}(\tau) = R_{yy}(\tau) \quad (19)$$

Assuming the time delay $\tau = 0$ and using (15) and (19), the correlation coefficient function can be obtained:

$$r_{xy}(0) = \frac{C_{xy}(0)}{\sqrt{C_{xx}(0)C_{yy}(0)}} \quad (20)$$

where r_{xy} is the correlation coefficient between signal $\{x(t)\}$ and $\{y(t)\}$ in time domain, which ranges from -1 (complete linear inverse correlation) to 1 (complete linear correlation) with $r_{xy} = 0$ meaning lack of linear interdependence. The sign of r_{xy} indicates the direction of correlation: $r_{xy} < 0$ implies inverse correlation, i.e., the two power signals are out of phase or oscillate with each other while $r_{xy} > 0$ implies direct correlation, i.e., a tendency of both signals to have similar value with the same sign.

The CCI between the i^{th} input signal x_i and the j^{th} output signal y_j is defined as:

$$r_{ij} = \frac{C(x_i, y_j)}{\sqrt{C(x_i, x_i)C(y_j, y_j)}} \quad (21)$$

where i and j are the input and output signal number, respectively.

The CCI will be a criterion to find the strong correlated measurements of different locations as inputs of the estimation model.

3.2 Input location selection algorithm for ARX model construction

The ranking of input location correlations are obtained by calculating CCI using (21) with measured signals in different locations. The performance of the ARX model is the criterion to evaluate the selected input signals. Please note that the ‘‘fixed location’’ is the output location of ARX model, where the measurement is currently inaccessible. The target of the algorithm is to obtain an accurate ARX model for the dynamic response estimation. The whole algorithm is described in Fig. 1. The basic steps outlined by this flowchart are:

- Step 1. Detrend the candidate measured signals. In order to obtain a wide-sense stationary random process, all measurement data should be detrended to remove direct-current components. Here a simple zero-mean filter should be used.
- Step 2. Use the detrended signals to calculate the CCI between fixed location and the candidate input locations.
- Step 3. Choose several inputs with the highest correlation coefficients as the ARX model inputs. For the EI system, only four inputs with the highest CCI are selected since the lower inputs the lower complexity of the ARX model. These four inputs are usually enough to get acceptable model accuracy in the following cases. However, different power systems can require a higher number of inputs, which should be further checked.
- Step 4. Train the ARX model with the selected input signals by CCI. The generation trip is considered as the excitation source to train the model.

Step 5. Test the ARX model accuracy. The model accuracy can be reflected by the accuracy index. ARX models with low accuracy index will be rejected while only the model with the acceptable accuracy can be selected as the final model for further estimation studies. The threshold of the accuracy index can be defined according to the study system and the event type [20]. The threshold or EI simulation case study is 0.9.

Usually, an ARX model can be obtained with acceptable accuracy e , i.e., model accuracy higher than preset accuracy threshold, ε , in one or several iterations. However, to prevent the proposed algorithm from infinite iteration, an outer loop is used to make sure the algorithm can successfully exit within limited iterations, i.e., maximum iteration number of N_{max} .

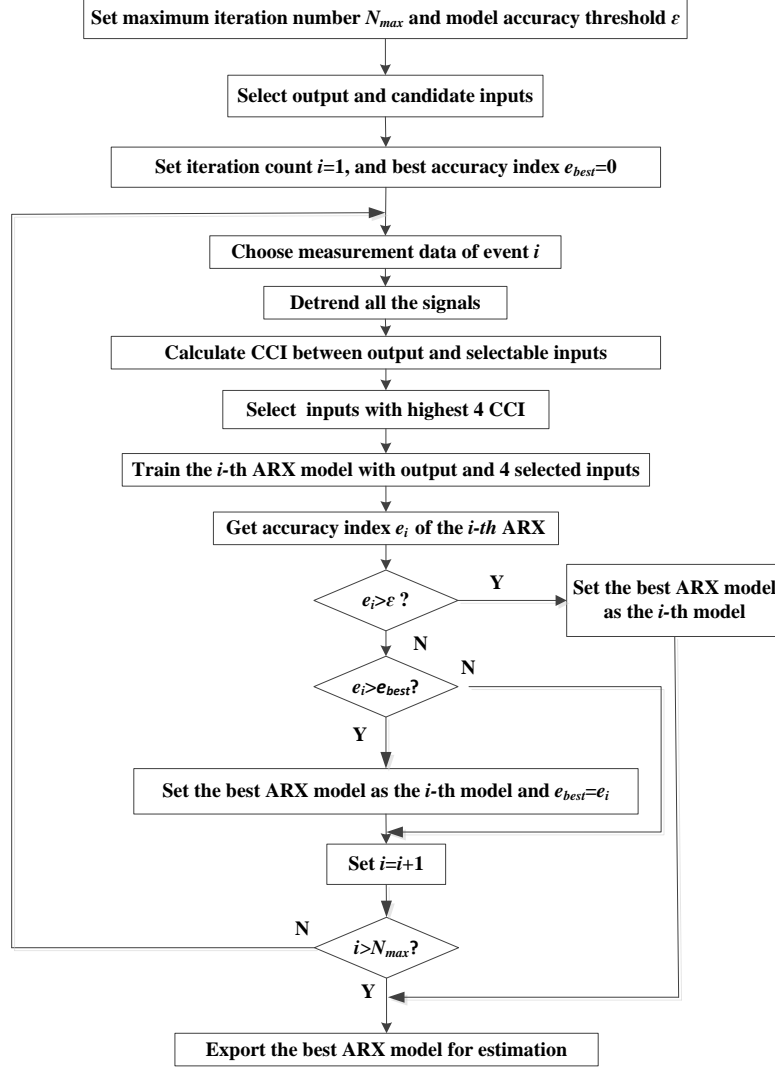


Fig. 1. Flowchart of the input selection approach for model construction

3.3 Correlation coefficient map

A correlation coefficient map (CCM) can be developed for large power systems with the MATLAB® mapping toolbox to show the CCI in a visual way. The EI system can be taken as an example to explain how to obtain a CCM. In the example shown in Fig. 2(a)-(c), the artificial measurement data were generated by simulating an 838-MW generation trip in eastern Alabama. In the CCM, different colors are used to indicate the correlation coefficient values. The color bar on the right of the map shows the coefficient value. In the map, one measurement location is chosen as the fixed location, which is the CCI calculation reference location. Then, the correlation coefficients between this fixed location and the other measurement locations can be shown vividly on the map. In Fig. 2, the correlation coefficients of the measured voltage magnitude, phase angle, and frequency between the fixed location and other locations can be found from the map, respectively. If the fixed location is the ARX model output location, the measured locations sharing a similar color can be selected as the input locations. From the CCM in Fig. 2, the frequency and phase angle correlation are strong for the entire system while the voltage magnitude correlation is shown as weak over wide areas. In this paper, only frequency and phase angle dynamic response estimations are studied.

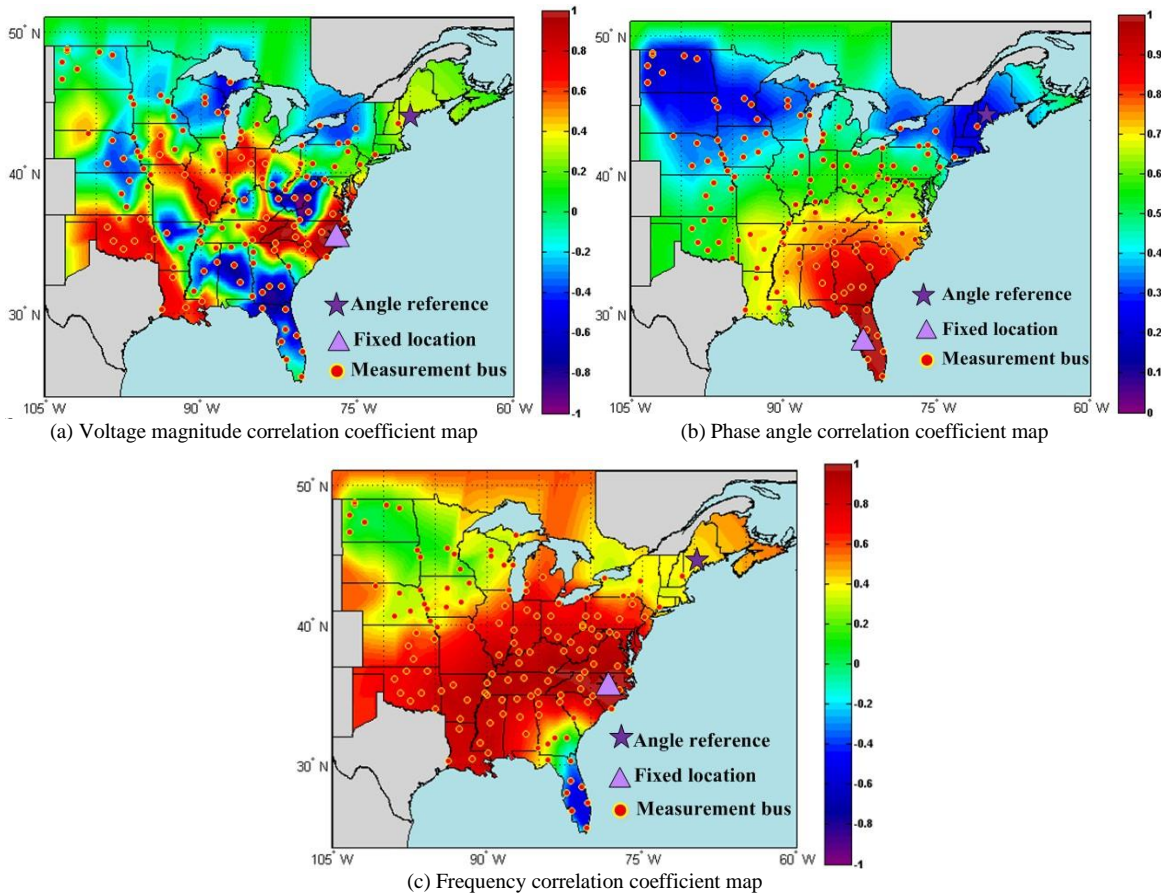


Fig. 2. Correlation coefficient map

4. Case Study

To demonstrate the effectiveness of the proposed approach in dynamic response estimation, the simulation was used to simulate the artificial measurements. A case study was performed on the EI system, which contains 3,000 generators and 16,000 buses in the PSS/E EI model. 135 buses are assumed to be measured by PMUs, and these 135 buses are evenly distributed throughout the EI system. The simulation results for the 135 buses are considered as the measured signals and 20 seconds of data after one disturbance is used. Note that the EI system serves as the test system in following cases. Two events were simulated to generate the artificial measurement data. Event 1 was an 814-MW generation trip in southeastern Florida, and Event 2 was an 838-MW generation trip in eastern Alabama, both of which are shown in Fig. 3 with blue stars. Event 1 was used to train the ARX model and Event 2 was used to perform the verification of the proposed algorithm. CCM were obtained from the frequency and phase angle response during Event 1, respectively, and they are shown in Fig. 3. Based on the CCM information, various sets of inputs were selected, which were then used to estimate the dynamic response during Event 2 at the fixed location (shown in Fig. 4 and Fig.5). A comparison of actual versus estimated responses and the accuracy index are provided in Fig.4.

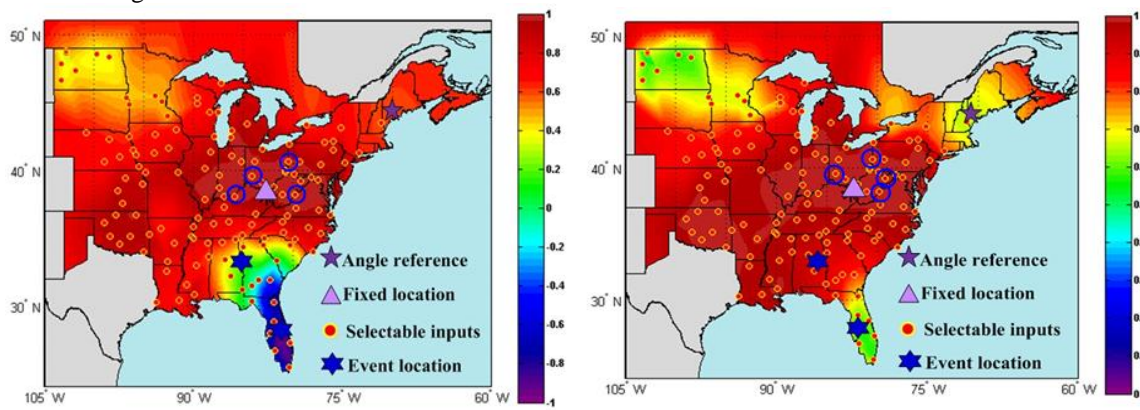
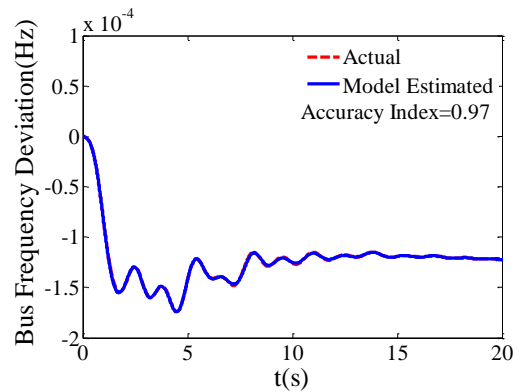
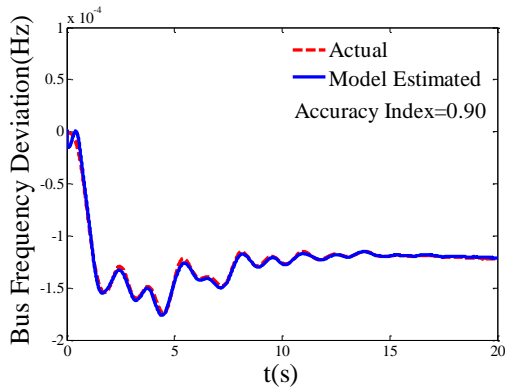
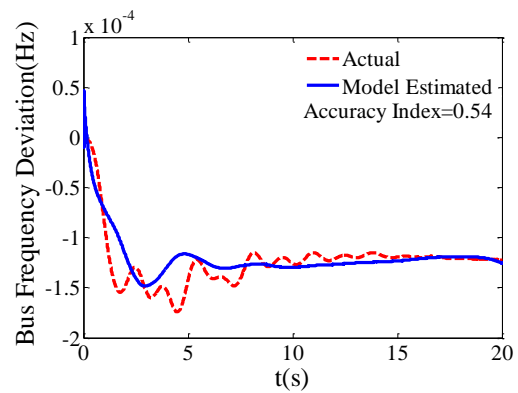
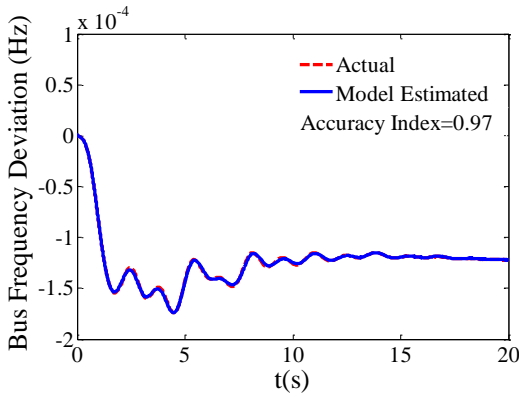


Fig. 3 (a). Frequency correlation coefficient map

Fig. 3 (b) Phase angle correlation coefficient map



(a) Highest 4 and lowest 4 coefficient locations as ARX model input locations (b) Highest 4 coefficient locations as ARX model input locations

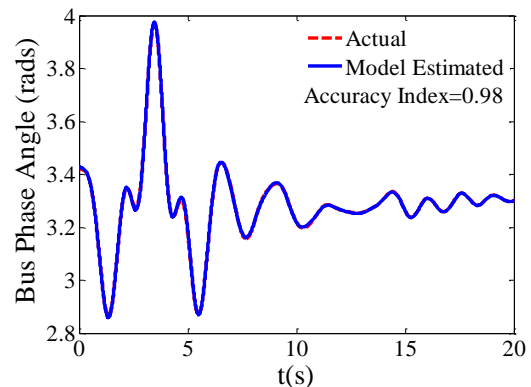
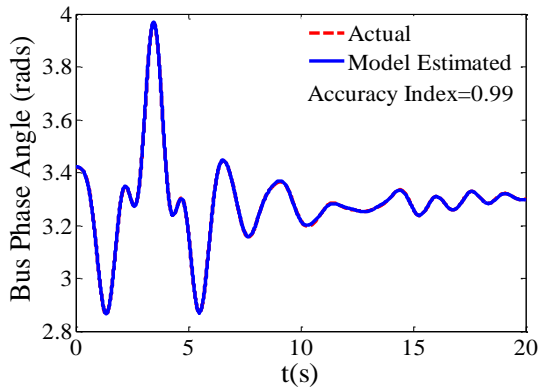


(c) Highest 1 coefficient location as ARX model input location

(d) Lowest 1 coefficient location as ARX model input location

Fig. 4. Input location selection algorithm verification using frequency dynamic response estimation

From the comparison results between the actual and estimated frequency responses in Fig. 4, it is obvious that the estimated frequency responses match the actual ones very well when input locations have high correlation coefficients, as shown in Fig. 4 (a)-(c) for frequency. If input locations with very low correlation coefficients are selected, the estimated responses do not match the actual response and the model estimation accuracy index is very low.



(a) Highest 4 and lowest 4 coefficient locations as ARX model input locations (b) Highest 4 coefficient locations as ARX model input locations

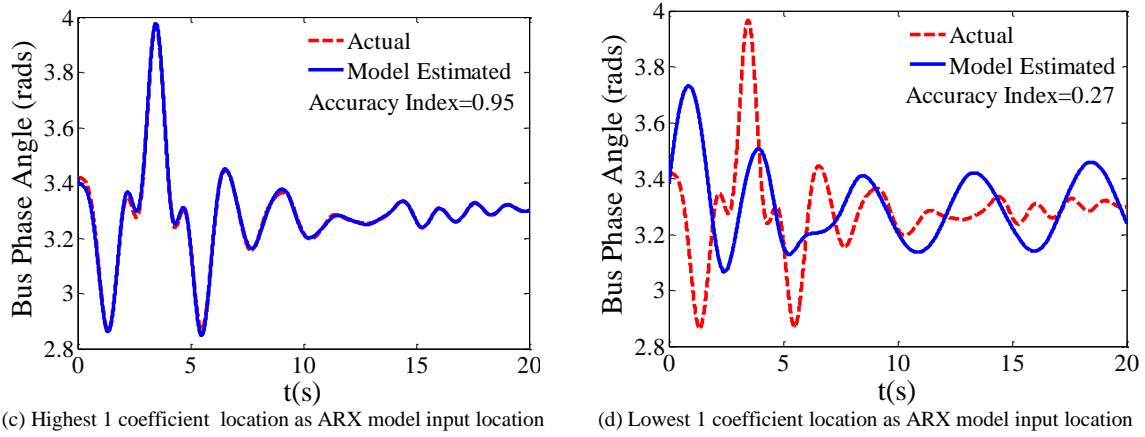


Fig. 5. Input location selection algorithm verification using phase angle dynamic response estimation

The comparison results between the actual and estimated phase angle responses are given in Fig. 5. It is obvious that the model estimated responses match the actual ones very well when input locations have high correlation coefficients, as shown in Fig.5 (a)-(c). On the contrary, the estimation result may be completely wrong like the estimation result in Fig.5 (d).

From these two dynamic response estimation results in Fig. 4 and Fig. 5, it can be seen that only the locations with higher correlation coefficients should be selected for the model estimation. Otherwise, the estimation results will be very bad. According to the input selection results, the selected highest four input locations are shown in Fig.3 with blue circles.

5. Correlation Coefficient Index Robustness Test

The results from the previous section show that the proposed input location selection approach can help achieve the accurate estimations of frequency and phase angle in the test cases. However, the proposed correlation calculation method is based on the disturbance data, which means that the CCI may be influenced by the disturbance location and magnitude. Therefore, the robustness of this approach still needs to be verified in order to be applied to a real power system.

5.1 Disturbances in the same location with different magnitudes

The test system is the same as in Section 4. The event location is shown in Fig. 6 with a red star. In the PSS/E dynamic simulation, the five event locations are the same while the generation trip magnitudes are different. Table 1 shows the event details. The test results are shown in Fig.7(a) and Fig.7(b), respectively.

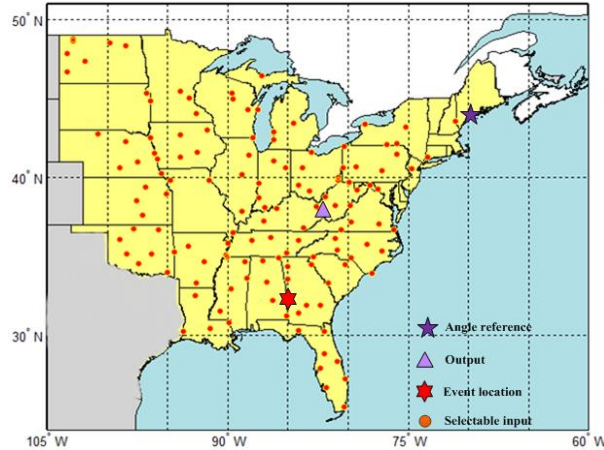


Fig. 6. Event location and details of the test system

Table 1 Generation trip events for the robustness test

Event Number	1	2	3	4	5
Amount (MW)	328	510	717	863	1,149

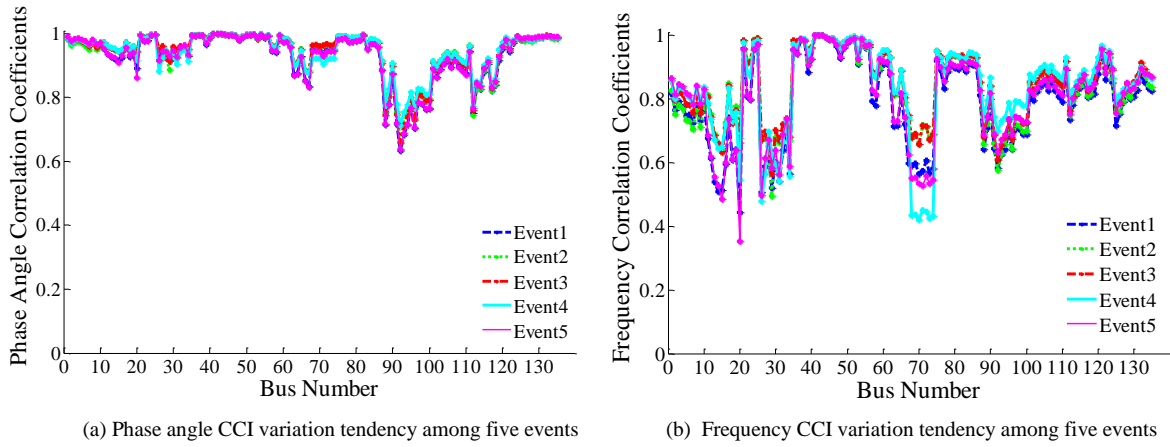


Fig. 7 CCI variation tendency among five events

As shown in Fig. 7, for both phase angle and frequency, the CCI variation tendencies remain similar for all the events, which indicate that the magnitudes do not influence the correlation relationship much and the proposed approach is robust to different disturbance magnitudes.

5.2 Disturbances in different locations with similar magnitudes

As shown in Fig. 8 by the red star, five generation trip events, each of which is around 1,000-MW, occur independently in the south, north, west, east, and middle of the EI system (to show the universality of the event locations). The output location is labeled by the purple triangle in Fig. 8 and the event details are shown in Table 2. The phase angle and frequency CCI variation tendencies are shown in Fig. 9.

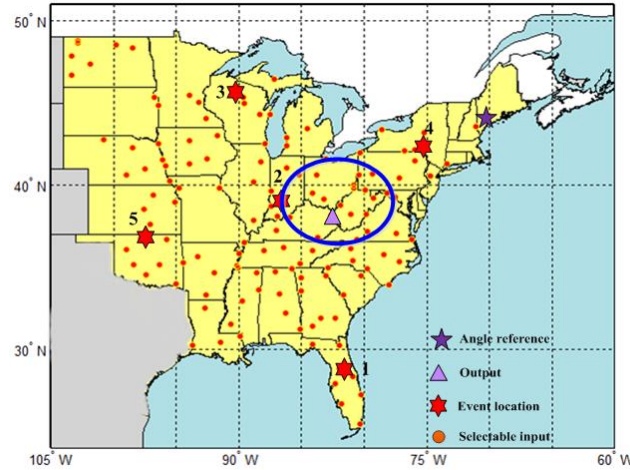


Fig. 8. Event locations and details of the test system

Table 2 Generation trip events for the robustness test

Event Number	1	2	3	4	5
Amount (MW)	957	1,000	1,017	1,150	990

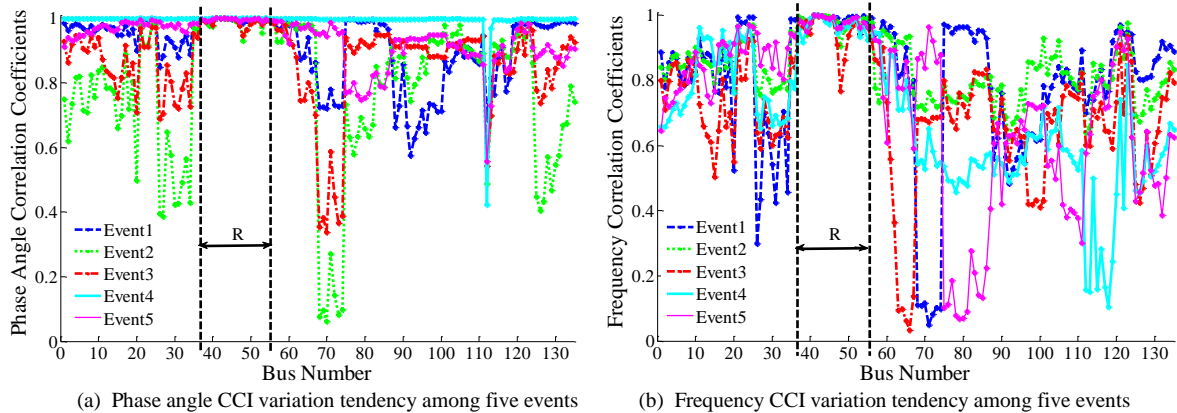


Fig. 9. CCI variation tendency among five events

Fig.9 indicate that most of the CCI values change when an event occurs in different locations. However, there are always some measurement locations which have very stable and high correlation coefficients no matter where the event occurs. These CCI variation trends are shown in Fig.9 with the part marked 'R'. For this part, the measurements are all in the blue circle in Fig.8. Additionally, it is apparent that all these measurement locations are close to the output location. Using the proposed input location selection approach, all of the four selected input locations are in this stable strong correlation area. That also means the ARX model input locations selected by the proposed approach always have strong and robust correlation with the ARX model output location.

5.3 Strong correlation areas

Subsections 5.1 and 5.2 prove that the strong correlation part of input locations has good robustness, and a strong correlation area exists around the output location. In this section, five typical output locations (purple triangles in Fig.10) are chosen to find their strong correlation areas, which are shown in the blue circles in Fig. 10, and the area diameters are shown in Table 3.

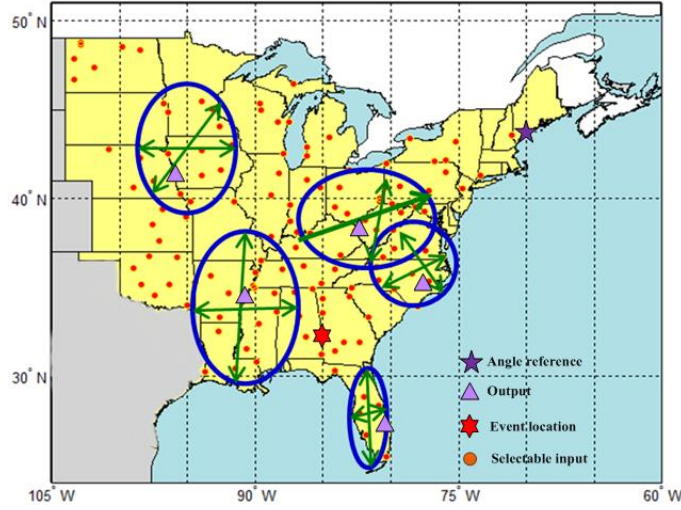


Fig.10. Strong correlation area for different outputs

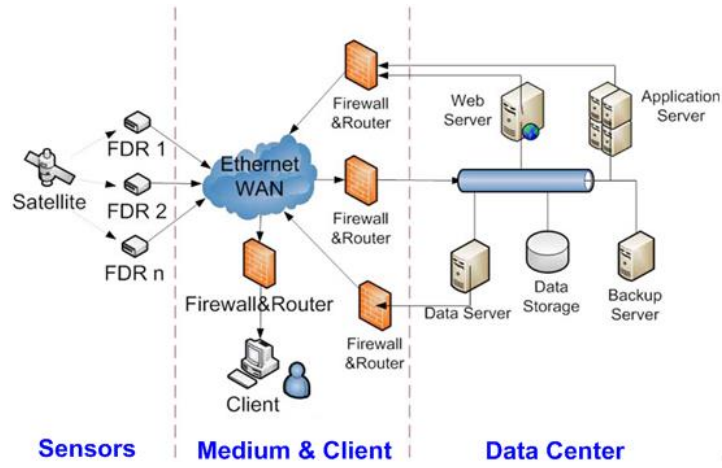
Table 3 Strong correlation area for different outputs

Interested Location	Florida	North Carolina	East Iowa	West Virginia	East Mississippi
Largest Diameter(mile)	320	240	390	520	600
Smallest Diameter(mile)	110	230	320	400	400

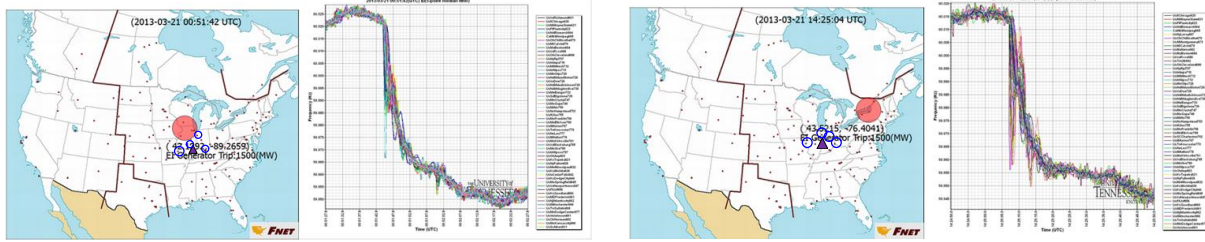
The results in Fig. 10 give a geographic area of the input location selection for dynamic response estimation. In other words, all the measured signals in the blue circle can represent inputs of the measurement-based model to estimate the dynamics at the location of the purple triangle. Therefore, the strong correlation areas will benefit the selection of input locations.

6. Algorithm Test Using Real Measurement Data

The power system frequency monitoring network (FNET) is a wide-area measurement system that takes high accuracy, GPS-synchronized measurements at standard end-user distribution voltages [26], which is shown in Fig. 11(a). As a member of the PMU family, the FDRs used in the FNET system measure frequency, voltage and phase angle at standard 120V outlets and transmit these measurements through the Internet [27]. It serves the entire North American power grid through advanced situational awareness techniques, such as real-time event alerts, accurate event location estimation, animated event visualization, and post event analysis [28]. The FDR measurements used here are the frequency and phase angle signals sampled at 0.1 seconds. As shown in Fig. 11(b) and Fig. 11(c), two disturbances monitored by FDRs were selected in this test. The red circles are the detected event locations and the purple triangle is the ARX output location. Event 1 is used to calculate the CCI and train the ARX model using the proposed input location selection approach. The estimation results and the estimation accuracy index are shown in Fig. 12.



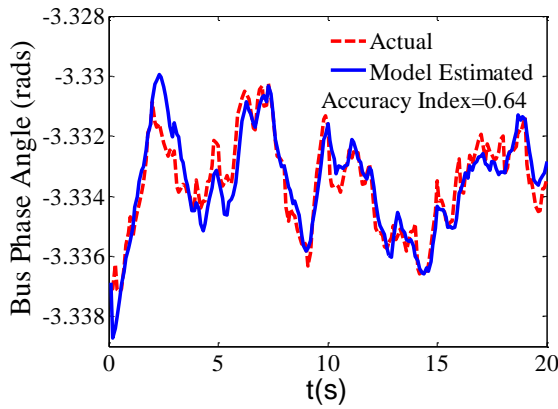
(a) Architecture of the FNET system at University of Tennessee, Knoxville, USA



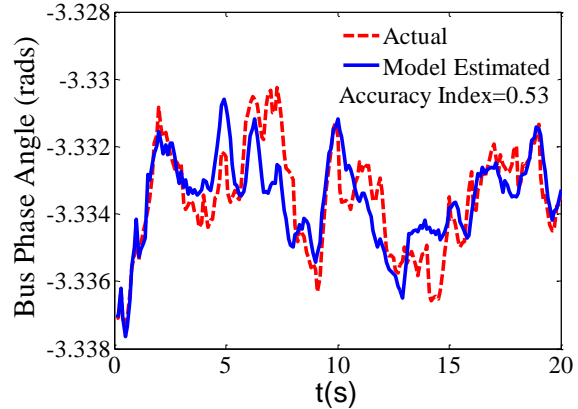
(b) Event 1 and measured frequency responses by FDR

(c) Event 2 and measured frequency responses by FDR

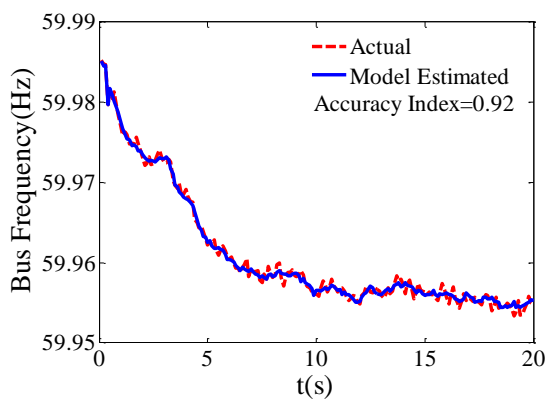
Fig. 11 FNET and display of frequency response measured by FDR



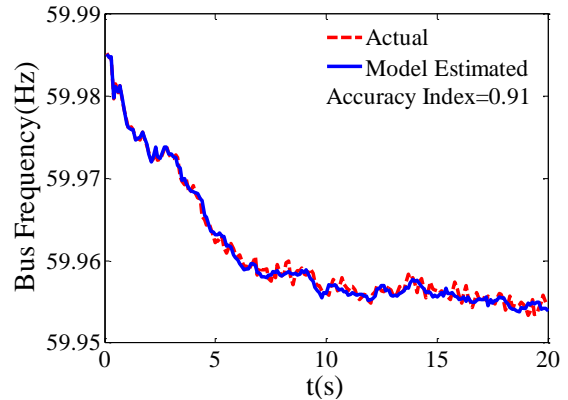
(a) Highest 4 coefficient locations as ARX model input locations



(b) Highest 1 coefficient location as ARX model input location



(c) Highest 4 coefficient locations as ARX model input locations



(d) Highest 1 coefficient location as ARX model input location

Fig. 12. Bus frequency and phase angle dynamic response estimation

As shown in Fig. 12, it demonstrates that, even one input location may be enough for the frequency to get an accurate estimation in some cases, while four input locations are needed for phase angle estimation. The highest four FDR locations are marked as blue circles in Fig. 11. Comparing the accuracy index in Fig.5, and Fig. 12, the estimation accuracy using real measurement data is lower than using simulation data. The main reason is that the real measurement data contains noise.

7. Conclusions

The proposed correlation coefficient index is an effective measurement-based approach to calculate the correlation relationship between different measurements at different locations. Moreover, the strong correlated input locations for dynamic response estimation can be quickly identified using the proposed approach. In this paper, the proposed algorithm is verified using both artificial measurement data generated by simulation and real measurements from FDRs. The complexity of the ARX model can be reduced consequently and the estimated results match the real dynamics very well. Furthermore, the real FDR measurement test indicated that the proposed approach gave a stable performance and can be implemented in the actual power systems for the missing data estimation. Besides, the developed correlation coefficient map could visually demonstrate the correlation between different measurement locations. In the future, ambient data will be considered to study the normal operation estimation.

Acknowledgement

This work is partially supported by the Electric Power Research Institute and also made use of Engineering Research Center Shared Facilities supported by the DOE under NSF Award Number EEC1041877. Additional support is provided by the CURENT Industry Partnership Program and China National Government Building High-level University Graduate Programs ([2012]3013). The authors also gratefully acknowledge the FNET team at the University of Tennessee for providing the phasor measurement data from FNET/Grideye for the verification work.

References

1. G. Phadke and J. S. Thorp, "Synchronized phasor measurements and their applications," Springer-Verlag, New York, NY, USA, 2008
2. Z. Zhong, C. Xu, B. J. Billian, L. Zhang, S.-J. S. Tsai, R. W. Conners, V.A. Centeno, A. G. Phadke, and Y. Liu, "Power system Frequency Monitoring Network (FNET) implementation," *IEEE Trans. on Power Systems*, 2005, 20,(4), pp. 1914-1921
3. K. Sun, S. Lee, P. Zhang, "An adaptive power system equivalent for real time estimation of stability margin using phase-plane trajectories," *IEEE Trans. on Power Systems*, 2011, 26, (2), pp. 915-923
4. A. G. Phadke and R. M. de Moraes, "The wide world of wide-area measurement," *IEEE Power Energy*, 2008, vol. 6, no.5, pp. 52-65
5. J. Dong, X. Ma, S. M. Djouadi, H. Li, Y. Liu, "Frequency prediction of the power systems in FNET based on state-space approach and uncertain basis functions," *IEEE Trans. on Power Systems*, 2014, 26, (2) (Accepted).
6. K. Sun, Q. Zhou, Y. Liu, "A Phase Locked Loop-based Approach to Real-time Modal Analysis on Synchrophasor Measurements", *IEEE Trans. Smart Grid*, 2014, 5, (1), pp. 260-269.
7. I. Kamwa, R. Grondin, Y. Hebert, "Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach," *IEEE Trans. on Power Systems*, 2001,16, (1),pp. 136-153.
8. R. Diao, K. Sun, V. Vittal, R. J. O'Keefe, M. R. Richardson, N. Bhatt, D. Stradford, S.K. Sarawgi, "Decision Tree-Based Online Voltage Security Assessment Using PMU Measurements", *IEEE Trans. Power Systems*, 2009, 24,(2), pp.832-839.
9. Y. Huang, M. Esmalifalak, H. Nguyen, R. Zheng, Z. Han, H. Li, L. Song, "Bad data injection in smart grid: attack and defense mechanisms," *IEEE Communications Magazine*, 2013,51, (1), 2013, pp.27-33.
10. J. Kim, L. Tong, "On topology attack of a smart grid: undetectable attacks and countermeasures," *IEEE Journal on Selected Area in Communications*, 2013, 31,(7), 2013, pp.1294-1305.
11. P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, "Modeless Reconstruction of missing synchrophasor measurements," Accepted by 2014 IEEE PES General Meeting.
12. V. Miranda, J. Krstulovic, H. Keko, C. Moreira, J. Pereira, "Reconstructing missing data in state estimation with autoencoders," *IEEE Trans. Power Systems*, 2012,27,(2), pp.604-611.
13. J. Qi, K. Sun, W. Kang, "Optimal PMU Placement for Power System Dynamic State Estimation by Using Empirical Observability Gramian", accepted by *IEEE Trans. on Power Systems*.
14. F. Aminifar, M. Shahidehpour, M. Fotuhi-Firuzabad, S. Kamaliniya, "Power system dynamic state estimation with synchronized phasor measurements," *IEEE Trans. on Power Systems*, 2014, 63, (2), pp.352-363.
15. E. Ghahremani, I. Kamwa, "Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements," *IEEE Trans. Power Systems*, 2011, 26, (4), pp.2556-2566.
16. I. Kamwa, B. Baraboi, R. Wamkeue, "Sensorless ANN-based speed estimation of synchronous generator: improved performance through physically motivated pre-filters," *IJCNN*, Canada, 2006, pp.1710-1718.
17. A. D. Angel, P. Geurts, D. Ernst, M. Glavic, L. Wehenkel, "Estimation of rotor angles of synchronous machines using artificial neural networks and local PMU-based quantities." *ELSEVIER Neurocomput*, 2007, Vol.70, Issues 16-18, pp. 2668-2678.
18. L. Ljung, *System identification: Theory for the user* 2nd Ed, PTR Prentice, New Jersey, USA, 1999, pp. 79-88.
19. L. Ljung, *System identification toolbox™ user's guid*, Matlab R2014b. pp. 292.
20. N. Bhatt, "Synchrophasor-based power system dynamic modeling for stability estimation," EPRI, Palo Alto, CA, Tech. Rep. 2012.1026456.
21. Sindelar, R., Babuska, R., "Input selection for nonlinear regression models," *IEEE Trans. on Fuzzy Systems*, 2004, 12, (5), pp. 688-696

22. A. Poncet and G. S. Moschytz, "Optimal order for signal and system modeling," *IEEE Int. Symp. Circuits and Systems*, 1994, pp.221-224.
23. Y. Liu, K. Sun, Y. Liu, "Measurement-based power system dynamic model for response estimation," *IEEE PES Power and Energy Society General Meeting*, San Diego, USA, 2012.
24. D. J. Trudnowski, "Estimating electromechanical mode shape from synchrophasor measurements," *IEEE Trans. Power Systems*, 2008,23, (3), pp.1188-1195.
25. K. Sun, K. Hur, P. Zhang, "A new unified scheme for controlled power system separation using synchronized phasor measurements", *IEEE Trans. Power Systems*, 2011,26, (3), pp.1544-1554.
26. J. Dong, T. Xia, Y. Zhang, Tony Weeks, James S. Thorp, and Yilu Liu, " Monitoring power system disturbances at the distribution level," *IEEE PES. General Meeting*, Pittsburgh, Pennsylvania, USA, 2008.
27. Y. Ye, R. Matthew Gardner, and Y. Liu, "Oscillation analysis in western interconnection using distribution-level phasor measurements," *IEEE PES ISGT*, Washington, D.C., USA, 2012.
28. Y. Zhang, P. Markham, T. Xia, et al., "Wide-area frequency monitoring network(FNET) architecture and applications, " *IEEE Trans. on Smart Grid*, 2010, 1,(2), pp. 159-167.